In presenting the dissertation as a partial fulfillment of the requirements for an advanced degree from the Georgia Institute of Technology, I agree that the Library of the Institute shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to copy from, or to publish from, this dissertation may be granted by the professor under whose direction it was written, or, in his absence, by the Dean of the Graduate Division when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this dissertation which involves potential financial gain will not be allowed without written permission.

3/17/65
b

THE OUTPUT OF QUEUEING SYSTEMS

A THESIS

Presented to

The Faculty of the Graduate Division

by

Peter Edward Chesbrough

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Industrial Engineering

Georgia Institute of Technology

June, 1966

THE OUTPUT OF QUEUEING SYSTEMS

Approved:

Chairman

Date approved by Chairman: 6/1/66

ACKNOWLEDGMENTS

I would like to express my appreciation to Drs. B. C. Spradlin, A. L. Hanken, and P. B. Han for their service on the Reading Committee. Special thanks is due to Dr. David E. Fyffe who acted as thesis advisor.

In a more general sense, I would like to thank all parties and institutions who have contributed to the support of my graduate work and the completion of this thesis.

I owe special acknowledgment to Mrs. Betty R. Sims who did an excellent job in typing the text and figures of a difficult manuscript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

GLOSSARY OF SYMBOLS

a	Number of arrivals
$a(t_{ba})$	A general independent interarrival time density function
$\alpha_x(t)$	The statement, "x arrivals occur in time t"
A_n	Ratio p_n/p_o for $M_n\|M_n\|1$ models
$\alpha(t)$	A general independent density function for waiting time until the next arrival
b	Number of departures (service completions)
$b(t_s)$	A general independent service time density function
$\beta(t)$	A general independent density function for waiting time until a service which was in progress at $t = 0$ is completed
c	Number of channels
$d_x(t)$	The statement, "x departures occur in time t"
D	Refers to a constant distribution
$\delta_{i,j}$	The Kronecker delta of order 2
e	As a superscript, indicates that the function is in the Laplace transform domain
E_k	Refers to the k^{th} Erlangian distribution
$E(\ )$	Expected value
$E(\ ,\ ,\ )$	Event
$\eta_o(t)$	A general independent distribution for the probability that waiting time to the next arrival exceeds t.
$f_n(t)$	The transient portion of $P_n(t)$
g	As a superscript, indicates that the function is in the geometric transform domain
$g(t_{ba})$	Negative exponential interarrival time density function

$g(t_{ba}\|\lambda_n)$	Negative exponential interarrival time density function when arrival rate is λ_n
G	Refers to a general distribution
GI	Refers to a general distribution with independence assumptions
$\gamma_o(t)$	A general independent distribution for the probability that a service which was in progress at $t = 0$ has not terminated by time t
$h(t_{bd})$	Density function for time between departures
i,j,k,m	Summation variables
ℓ	Number of input sources
λ	Mean arrival rate
λ_n	Mean arrival rate when n units are in the system
$\mathcal{L}(\ )$	Laplace transform
$\mathcal{L}^{-1}(\ )$	Inverse Laplace transform
M	Refers to a Poisson occurrence distribution or to a negative exponential inter-event time distribution
M_n	Refers to a state-dependent M distribution
μ	Mean service rate
μ_n	Mean service rate when n units are in the system
n	State of (number of units in) the system; the total units in the queue and the service facilities
N	Total number of units allowed in a truncated queueing system; occasionally (e.g. in the literature on tandem queues), the size of the interstage bank
$\nu_o(t)$	A general independent distribution for the probability that service time on a unit exceeds t
P_n	Steady-state probability of n units in the system
$P_{\geq n}$	Steady-state probability of n or more units in the system
$P_n(t)$	Probability that there are n units in the system at time t
$\Phi_a(t)$	Probability of a (Poisson) arrivals in time t

$\Phi_a(t\|\sim)$	Probability of a (Poisson) arrivals in time t, given the partial information (~)
$\Psi_b(t)$	Probability of b (Poisson) service completions in time t
$\Psi_b(t\|\sim)$	Probability of b (Poisson) service completions in time t, given the partial information (~)
r_n	Probability that the state of the truncated M|M|c system is n at a departure epoch
$R_{\geq n}$	Probability of n or more units in the truncated M|M|c system at a departure epoch
s	The Laplace transform variable
$S_n(t)$	The statement, "n units in the system at time t"
t_{ba}	Time between arrivals
t_{bd}	Time between departures
t_s	Service time
$\theta_o(t)$	A general independent distribution for the probability that service time on a unit exceeds t
$v(t_s)$	Negative exponential service time density function
$v(t_s\|\mu_n)$	Negative exponential service time density function when the mean service rate is μ_n
z	The geometric transform variable
*	The convolution operator

SUMMARY

The thesis is concerned with the development of a theory of queueing output behavior for the important applications to queueing network analysis. The method of attack involves the stochastic description of an arbitrary interdeparture event as the union of a set of mutually exclusive and collectively exhaustive sub-events. Sub-event density functions for the time between departures are summed, using standard combinatorial probability theorems, to obtain the density function for the union. The method is used to find output distributions for the M|M|c, truncated M|M|c, and multichannel queues-with-discouragement models. Finally, an expression is derived for the output of the general multichannel model with state-dependent Poisson input and state-dependent negative exponential service time distributions.

A variation on this formulative technique is used to obtain the output distribution of the more general GI|GI|c model. An extention to the heterogeneous multi-input, heterogeneous multichannel model with general independent interarrival and service time distributions is described.

Theorems and results due to Burke (1956), Reich (1957), Finch (1958, 1959), and Chang (1963) are discussed in terms of their validity under relaxed restrictions, and their applicability to more general models.

CHAPTER I

INTRODUCTION

Networks of Queueing Systems

In recent years, with the advance of increasingly more complex industrial and other processes, a great deal of attention has been focused on systems analysis. One area of particular interest is that system in which units arrive at a service location, possibly wait for earlier arrivals to complete their services, are served, leave and proceed to a second service location to repeat the steps, and so on until a multi-step process has been completed. Examples of such networks may be found in department stores, cafeterias, telephone line systems, traffic flow, railroad switching yards, factory assembly lines, mail order houses, and many other everyday situations.

Study of these networks is usually initiated by analyzing them into discrete stages, each consisting of a service station and its associated waiting line. Inclusion of the waiting line permits one to take the output of one stage as the input to the next without having to perform a separate study of each inter-stage flow process.* Each stage is thus an isolated "queueing system," or "queue," of the form shown in

---

*When only a limited amount of inter-stage waiting space is available, there is a possibility of "blockage"(i.e. the first stage may become inoperative since a unit which has just completed service cannot vacate the service station if the inter-stage queue is filled to capacity). In this case, the two stages must be studied jointly.

Figure 1. By convention, c identical, but independent, service stations in parallel with a common input are regarded as comprising a single "multichannel" queueing system. Thus, in its most general form, a network of queues may be thought of as a combination of multichannel models in parallel (Figure 2), "tandem" (Figure 3), or feedback (Figure 4) arrays.

In the past, isolated queues have been quite extensively investigated and results have been obtained in a general form for most parameters and statistics of interest.* However, it is only recently that significant progress has been made in the study of networks of queues. The existing network literature, with a few exceptions, dates from Burke's result [8] of 1956: The output of the multichannel queue with Poisson input distribution for the number of arrivals in time t (or equivalently, negative-exponentially distributed interarrival times) and negative-exponentially distributed service time distributions is itself Poisson. Reich [64, 65] and Finch [25, 27] obtained partial converses to Burke's result (for the single channel model) and Chang [10] gave a method whereby, given certain information, one can determine the output of the single channel model with arbitrarily and independently distributed interarrival and service time distributions. Beyond this, little has been presented except applications of the Burke theorem and specific examples which give partial converses to the theorem. This network literature will be described in greater detail in Chapter II.

---

*Saaty [69] reviews many of the more important papers and makes available a bibliography of some 900 listings for more detailed study. See also Doig [20] and Lunger [51].

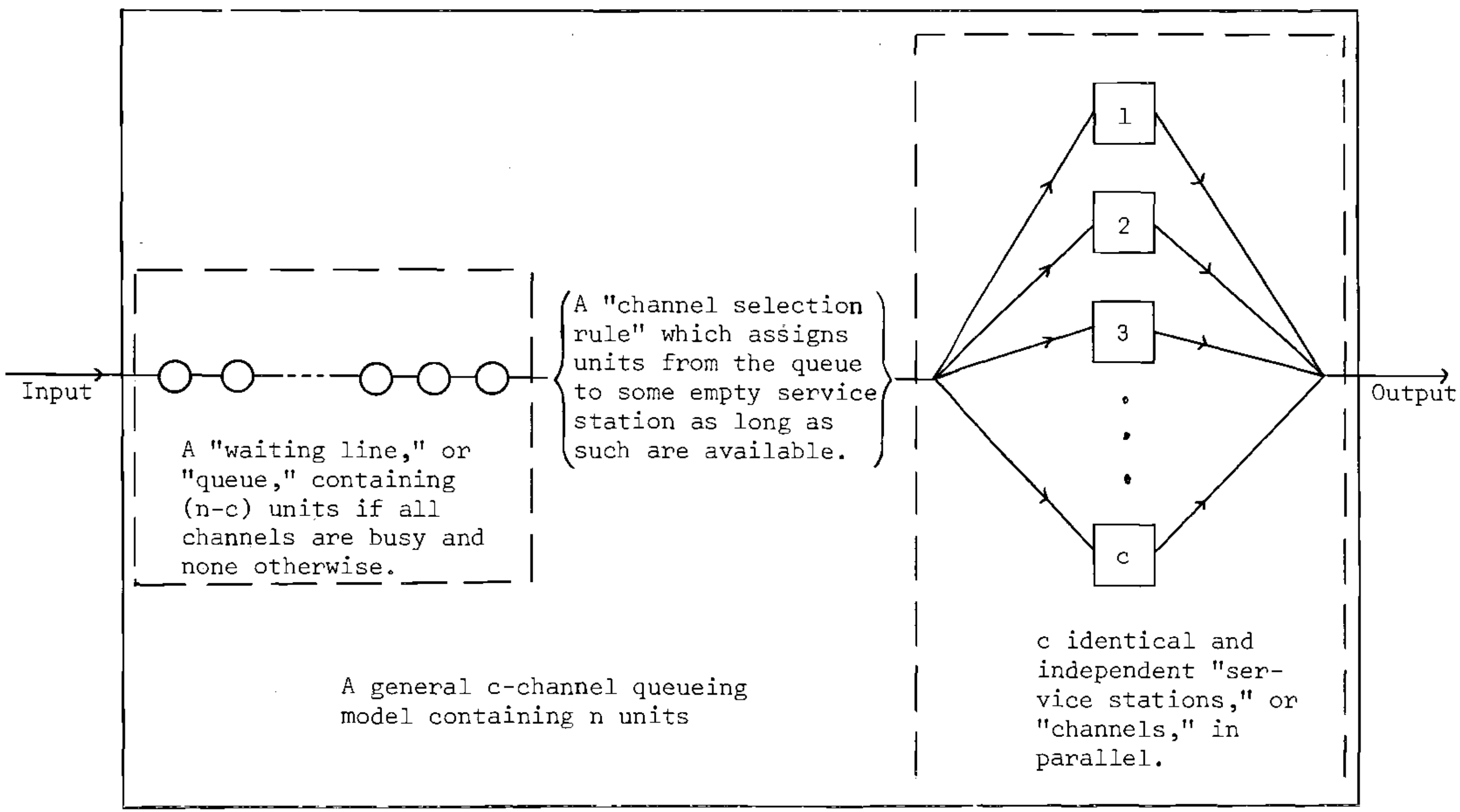


Figure 1. The General Multichannel Queueing System.

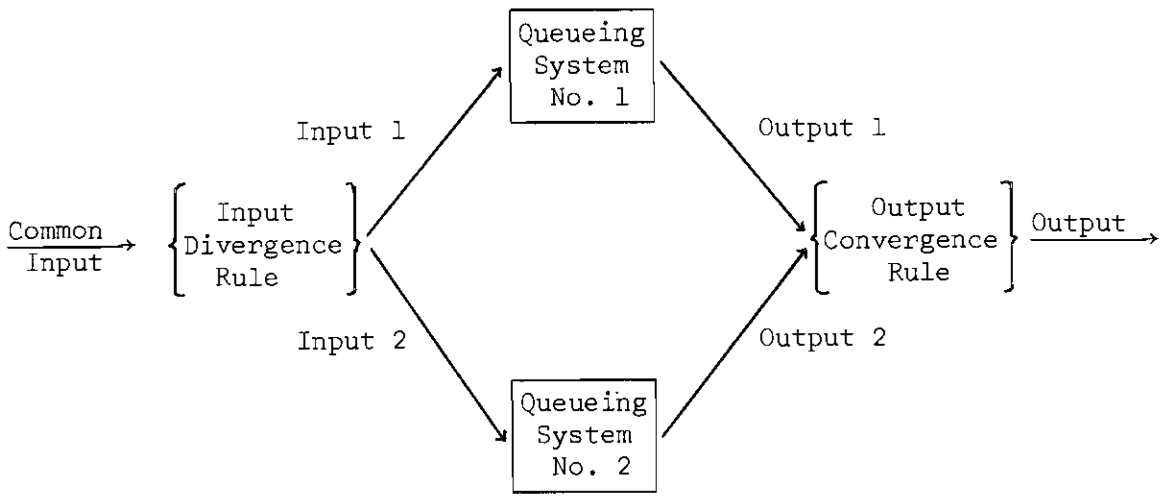


Figure 2. Two Queues in Parallel.

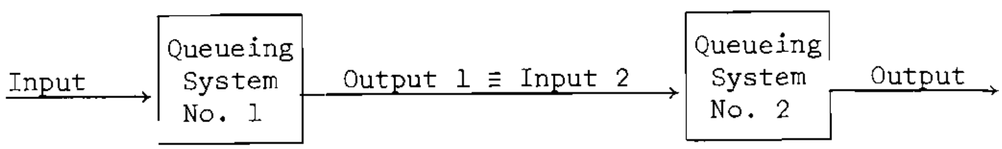


Figure 3. Two Queues in Series ("Tandem" Queues).

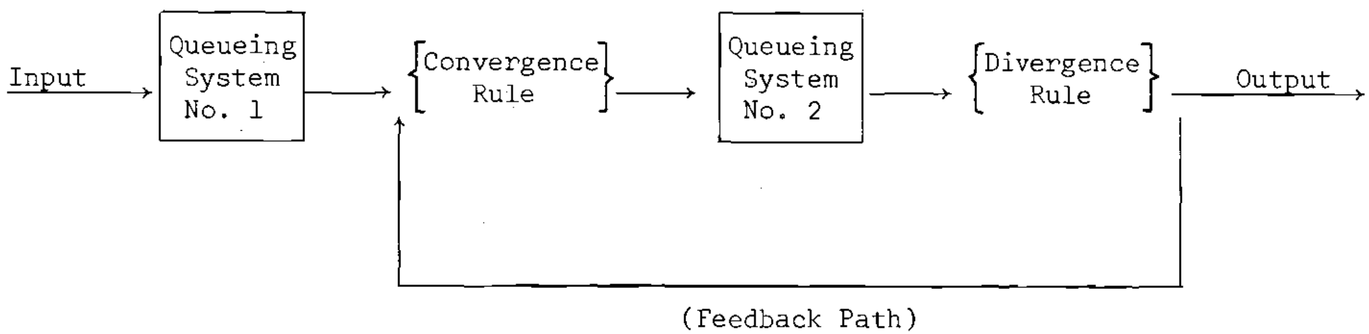


Figure 4. Example of Feedback in a Network of Queues.

Nature of the Present Investigation

Objective and Purpose

It is evident from the above remarks that the missing key to a thorough analysis of network queueing problems is a knowledge of the input to, or equivalently, the output of each queue in the network. The primary objective of the present paper is a derivation of output distributions for some common queueing models and a specification of general methods whereby output distributions might be obtained for other models. Secondary objectives are to provide output information in a form suitable for analysis of transient processes and processes with partial information and to ascertain the applicability of the Reich and Finch converses to the multichannel case. The purpose of the study is, of course, to obtain results which are applicable to queueing network analysis.

Method of Attack

The procedure used to obtain output distributions for multichannel models with state-dependent* Poisson input and state-dependent negative-exponentially distributed service times is a case enumeration and subsequent summation of corresponding probabilities of the possible interdeparture events. Standard combinatorial probability theorems are used in the summation to obtain the probability density function of the union of events. The interarrival times, service times, and interdeparture times are related by a simple algebraic equation for each possible set

---

*The state of a queueing system is defined as the number of units in the entire system, that is, the sum of those in the waiting line and those being serviced (See Figure 1).

of events. The convolution theorem for the sum of independent random variables is then used to obtain the density functions for the interdeparture time intervals.

A variation of this procedure is used to study more general models.

Form of Results

An exact equation is given for those models which were completely specified. A formula, a method, or both is given (or, in some cases, suggested) for the more general models.

Scope and Limitations

The results and the method of attack are applicable to most of the queues commonly encountered in the literature. However, unless some device can be found to examine dependent interdeparture events in a general form, the method cannot be extended to the model with dependent interarrival and service time distributions. Further, the result for the multichannel model with general independent interarrival and service time distributions is limited for application to those models for which the equilibrium state probabilities are known.

Assumptions

In all of the models treated, service will be assumed to be first-in-first-out. This condition may be somewhat relaxed for the models of Chapter III (e.g. last-in-first-out is acceptable); however, it is definitely necessary to the more general models of Chapter IV. No balking, reneging, feedback, or other complications are allowed except as interpretations of the mathematical formulation.

Kendall's Classification System

An effort has been made to retain a consistent, standard notation throughout this paper. The pertinent symbols are listed and defined in a glossary preceding this chapter. In addition, Kendall's abbreviated classification system for queueing models has been adopted. Kendall's system [43] uses M to denote a Poisson occurrence distribution or, equivalently, a negative exponential inter-event interval distribution; D for a constant distribution; E_k for the k^{th} Erlangian distribution; G for a general distribution; GI for a general distribution with independence assumptions; etc. The form $G|M|c$, for example, then denotes the c channel model with general input and negative exponential service time distribution. It has been necessary to make one addition to the Kendall system: In Chapter III, a Poisson distribution with state-dependent mean is denoted by the subscripted symbol M_n.

CHAPTER II

LITERATURE SURVEY

The relevant literature may be classified as dealing with the output of a queue or as dealing with networks of queues. The first group is pertinent to the primary objective of this paper (providing information useful for the analysis of queueing networks). Proofs (outlined) of the major results are included here for the sake of completeness. The second group, dealing primarily with specific examples of tandem queues, is based on the results of the first group or on other results which do not depend directly upon a knowledge of the output processes of the queues involved. While this literature is loosely relevant to the objectives of the paper, it is also rather extensive and we will be forced to briefly cite only a few representative papers. The purpose of including this secondary, peripheral material is to summarize past efforts at queueing network analysis and to suggest applications and extensions of the present work.

Output of a Queueing System

Burke [8], 1956, and later, Reich [65] and Finch [27] have independently established that the output of the steady-state, multi-channel queue with Poisson input and negative exponential service times is also Poisson. Burke's method involved the solution of a set of differential difference equations. Letting L denote the length of an arbitrary inter-departure interval and n(t) the state of the system at time t after the last previous departure, he wrote $F_k(t)$ for the joint probability that $n(t) = k$ and $L > t$. Using the Poisson arrival and service assumptions,* he obtained for an infinitesimal interval of length dt,

$$F_o(t+dt) = F_o(t)(1-\lambda dt),$$

within infinitesimals of higher order, since $L > t + dt$ if and only if $L > t$ and no arrival occurred during dt. Similarly, he obtained

$$F_k(t+dt) = F_k(t)(1-\lambda dt-j\mu dt) + F_{k-1}(t)\lambda dt \ ,$$

where $j = k$ for $k < c$ and $j = c$ for $k \geq c$.

In the limit as $dt \to 0$, these equations reduce to

$$F_o'(t) = -\lambda F_o(t)$$

and

$$F_k'(t) = \lambda F_{k-1}(t) - (\lambda+j\mu)F_k(t) \ ,$$

subject to the initial conditions (which imply the existence of equilibrium)

$$F_k(0) = p_k \ .$$

Burke then cited an inductive solution to yield

---

*These are listed in the preliminary remarks of Chapter III.

$$F_k(t) = p_k e^{-\lambda t} \tag{†}$$

as the unique solutions subject to the initial conditions. Thus the marginal distribution of the interdeparture intervals is negative exponential with parameter λ, the same as the distribution of the interarrival intervals.

The independence of L and n(L) was established as follows: The probability that $t + dt > L > t$ and $n(L+0) = k$ is

$$F_{k+1}(t)\ (k+1)\mu\ dt\ , \quad \text{for } k + 1 \leq c\ ,$$

and

$$F_{k+1}(t)\ c\mu\ dt\ , \qquad \text{for } k + 1 > c\ .$$

Upon substitution of the state probabilities, these expressions reduce to

$$\frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k p_o e^{-\lambda t}\lambda dt$$

and

$$\frac{1}{c!c^{k-c}}\left(\frac{\lambda}{\mu}\right)^k p_o e^{-\lambda t}\lambda dt\ ,$$

respectively. These latter expressions are factored into the marginal probability functions of n(L) and L, thus proving the independence of L and n(L). The mutual independence of all interdeparture intervals follows from the Markovian property of the negative exponential distribution. This completes Burke's proof.

A different method of proof which included Burke's result as a special case (of a Markov chain theorem) was developed by Reich [64, 65]. In the same paper, Reich also proved a partial converse to Burke's theorem: that, for a single channel queue, Poisson input and output implies either negative exponential service times or a step function at zero.* A third result was a proof, by contradiction, that the output process of the $E_2|E_2|1$ queue is not of type E_2. Thus we may not generally expect the output process of a queue to match the input, even in the steady state.

Finch [25, 27] showed that, for the single channel queue $M|G|1$,** toleration of an infinite queue and negative exponential servicing are necessary and sufficient conditions for the Poisson output result to be correct, for the independence of the inter-departure intervals, and for the independence of the queue length left by a departing unit. Finch's approach was essentially the same as Burke's. Achieving the equivalent of Equation (†) in terms of a general independent service distribution, he set the desired output and independence conditions and found that the service times must be negative-exponentially distributed.

The second major output result came in 1962 with Chang's method

---

*I suggest that Dr. Reich intended, as his second case, the implication of an impulse at $t = 0$ (i.e. the trivial case of zero service time). The case has not been discussed in the subsequent literature and I was unable to verify it with the methods of Chapters III and IV since I did not have an expression for p_o. Note that the step function result would be a contradiction to the Finch sufficiency condition given in the next paragraph.

**$M|G|1$ is a misnomer taken from the title of the Finch paper. Finch made independence assumptions on the distribution of service times which suggest the more proper designation, $M|GI|1$.

[10] for determining the output of the GI|GI|1 model. His derivation is as follows. Let τ_n, h_n be the waiting time and service times of the n^{th} item and let x_{n-1}, y_{n-1} be the interarrival time and the interdeparture time between the $(n-1)^{th}$ and the n^{th} items. Lindley [48] has given the waiting time of the n^{th} item in terms of the $(n-1)^{th}$ item as

$$\tau_n = \begin{cases} \tau_{n-1} + h_{n-1} - x_{n-1} \,, & \tau_{n-1} + h_{n-1} - x_{n-1} > 0 \,; \\ 0 \,, & \text{otherwise.} \end{cases}$$

Chang writes for the interdeparture time,

$$y_{n-1} = \begin{cases} h_n \,, & \tau_n > 0 \,; \\ h_n + u_n \,, & \tau_n = 0 \,, \end{cases}$$

where u_n denotes the length of time that the server has been idle before the arrival of the n^{th} item. It is evident that u_n can only be positive or zero. Letting u_n^* denote the positive value other than zero, Chang obtains

$$u_n = \begin{cases} u_n^* = x_{n-1} - h_{n-1} - \tau_{n-1} > 0 \,, & \tau_n = 0 \,; \\ 0 \,, & \tau_n > 0 \,, \end{cases}$$

and the simplified expression,

$$y_n = h_n + u_n \ . \qquad (\dagger)$$

The probability of u_n is treated as the sum of two functions.

$$p(u_n) = p(u_n^*) + p(u_n=0)$$

$$= p(u_n^*) + p(\tau_n>0)\delta(u_n) \ ,$$

where $\delta(u_n)$ is the unit impulse function,

$$\delta(u_n) = \begin{cases} 1 \ , & u_n = 0 \ ; \\ 0 \ , & u_n \neq 0 \ . \end{cases}$$

Let $P(s)$ be the Laplace transform of $p(u)$. Then, by the definition,

$$P(s) = E[e^{-su}] \ ,$$

and Cauchy's integral theorem, Chang obtains

$$E[e^{-su_n}] = \frac{-1}{2\pi j} \oint_c \frac{E[e^{-z(x_{n-1}-h_{n-1}-\tau_{n-1})}]}{z - s} dz + p(\tau_n>0) \ . \qquad (\ddagger)$$

The contour c is taken from $-j\infty$ to $j\infty$ and extends to the right half plane.

The representation up to this point is valid regardless of the dependence of the random variables involved. In the GI|GI|1 case at hand, x_{n-1}, h_{n-1}, and τ_{n-1} are independent random variables such that

$$E[e^{-z(x_{n-1}-h_{n-1}-\tau_{n-1})}] = E[e^{-zx_{n-1}}]\, E[e^{-zh_{n-1}}]\, E[e^{-z\tau_{n-1}}] \ .$$

Further, in the steady state, the distributions are identical for all items and the subscripts n and n-1 can be dropped in Equation (‡).

Assume that the Laplace transforms of the distributions exist. Let

$$F(s) = E[e^{-sx}] \ , \qquad B(s) = E[e^{-sh}] \ ,$$

$$D(s) = E[e^{-sy}] \ , \text{ and } \ W(s) = E[e^{-s\tau}] \ .$$

Note that $p(\tau=0)$ is just $1 - p_o$ where p_o is the probability that the system is empty when an item arrives. Equation (‡) becomes

$$P(s) = \frac{-1}{2\pi j} \oint_C \frac{F(z)B(-z)W(-z)}{z - s}\, dz + 1 - p_o \ .$$

From Equation (ǂ), it is known that the interdeparture time density function is the convolution of p(u) and the service time density function. Hence,

$$D(s) = B(s)P(s) = B(s)[\frac{-1}{2\pi j} \oint_C \frac{F(z)B(-z)W(-z)}{z - s}\, dz + 1 - p_o] \ .$$

Once D(s) is known, the interdeparture time distribution, d(y), can be determined from the inverse transform,

$$d(y) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} D(s)e^{sy}ds \ .$$

Chang's procedure is quite clear except for the means by which the waiting time density function, $w(\tau)$, may be obtained. He remarks that the waiting time distribution of this model may be obtained from Lindley's integral equation [48] of the Wiener-Hopf type, or from the contour integration used by E. Ventura [79]. In Chang's notation, Lindley's equation for the cumulative waiting time distribution, $W(\tau)$, is

$$W(\tau) = \int_{0}^{\infty} W(t)P'(\tau-t)dt \ ,$$

where
$$P'(\tau) = \int_{0}^{\infty} b(t)f(t-\tau)dt = \int_{0}^{\infty} b(t+\tau)f(t)dt \ ,$$

$$P(\tau) = 1 - \int_{0}^{\infty} b(t+\tau)F(t)dt \ ,$$

and t is the transformation variate $\tau - u$. For application, it is suggested that Lindley's equation can most easily be solved in the transform domain.

Networks of Queues

Tandem Queues

Most of the early analytical work with queues in series has been restricted to M|M|1 queues with phase-type servicing. Phase-type servicing refers to those tandem arrangements where there is no waiting between stages and, usually, where a unit must pass through all phases, in order, before his service is completed and another may begin. Although recent practice has been to use "phase" and "tandem" interchangeably, we will restrict our use of "phase" to its historical meaning. All models discussed will be assumed to be steady state, unless otherwise identified.

Perhaps the earliest example of phase-type servicing is the $M|E_k|1$ model. This model has been interpreted as k identical M|M|1 models in tandem with zero banks (no inter-stage waiting) and phase-type service (e.g., Sasieni [71], p. 145). If the exponential service times each have mean $1/(k\mu)$, then the system follows the k^{th} Erlang distribution with mean $1/\mu$.

Good [32], 1948, was among the first to study the number of individuals in a tandem system. O'Brian [62] treated the case of two M|M|1 models in series with an infinite inter-stage bank and gave expected queue lengths and expected waiting times. R. R. P. Jackson [39, 40] extended O'Brian's work to the case of k M|M|c models in tandem with infinite banks and different negative exponential service distributions for each stage. He obtained multivariate state probabilities, $P(n_1, n_2, \ldots, n_k)$, for the probability of various numbers of items of different stages, and multivariate waiting time distributions. J. R. Jackson [37, 38] further extended treatment to the case where Poisson arrivals were also allowed to enter any stage from outside the network.

Akaike [1] and Sacks [70] have studied the ergodic properties of two GI|G|1 queues in series, indicating an extension to several stages. Akaike required a customer to delay entering a second phase until a following customer entered the first phase. With this queue discipline, which guaranteed certain desired Markov properties, he found that ergodicity of the waiting times followed if $\lambda > \mu_1$ and $\lambda > \mu_2$. Sacks allowed infinite banks and showed that the condition for ergodicity of the waiting time distributions at the various stages is $\lambda > \max(\mu_1, \mu_2, \ldots \mu_k)$.

Nelson [61] developed an analytical method for calculating waiting times in networks of M|M|c queues with infinite banks. His work resulted from analyzing job shop production processes as part of the Management Science Research Project at the University of California. Nelson obtained the probability of waiting longer than a given time at all stages given the different exponential service times at each stage. DeBaun and Katz [18] have simplified Nelson's computations with a chi-square approximation to the sum of exponentials.

Luchak [50] has studied the M|GI|1 system in continuous time, indicating the application to phase-type servicing for appropriate choices of the service time distribution (e.g. $M|E_k|1$ in continuous time). Conolly [12, 13, 14] has applied a difference equation technique to good advantage in examining simple queues. In one paper [14], he studied queueing at a single serving point with group arrival. He showed that certain aspects of his model were equivalent to phase servicing of single arrivals.

Several recent papers have dealt with tandem networks with finite banks. The point of interest in these networks is the effect of blocking which occurs whenever a processed unit is unable to vacate a service facility because the bank ahead is already filled to capacity. Hunt [36] studied the utilization (or, "traffic density," defined as λ/μ) of M|M|1 models in tandem for several queue disciplines: (1) infinite banks before each stage, (2) finite banks at the second and succeeding stages, (3) a zero bank at each stage, and (4) zero banks and no vacant facilities (the entire line moves as one unit). Morris [60] dealt with the application of queues to materials handling analysis. He treated tandem queues and several other queueing network examples.

Makino [52, 53] presented two papers on the blocking effect in M|M|c tandem queues. In the first, he obtained a necessary condition for establishing expected queue size when N = 1 in a two-stage array of M|M|1 models. In the second, he studied the blocking effect in a two-stage M|M|1 array, a two-stage M|M|c array, and a three-stage M|M|1 array.

Other recent papers have treated non-Poisson queues and waiting time independence. Ghosal [31] has found the waiting time distribution for two-stage service, Poisson input, infinite inter-stage bank, and a gamma service time distribution at the first stage and an exponential service time distribution at the second. Suzuki [74] considered two queues in series with an infinite bank. The first is of type M|G|1 with service at the second given by $H_2(x) = 1 - e^{-\mu_2 x}$ $(x \geq 0)$. However, as Takàcs points out, "the author's results are incorrect because his proof is based on the false assumption that the queue sizes immediately before arrivals in the second queue form a Markov chain."*

---

*L. Takàcs, *Mathematical Reviews*, Vol. 29 (1965), No. 2873, p. 559.

Marshal and Reich [54] have presented a study of the characteristics of queues in tandem. Reich [66] has verified the independence of and obtained distributions for the inter-stage waiting times of $M|M|1$ models in series. Masterson, Gregory, and Sherman [55] studied an infinite sequence of $M|M|1$ models in series, concluding a negative answer to the conjecture, "some equilibrium statistics will be approached (with increasing n)." Loynes [49] has obtained necessary and sufficient conditions for a system of queues in series to be substable.

Burke [9] considered the dependence of delays in a two-stage $M|M|1$ array with infinite banks. He showed that the waiting times of the n^{th} customer at the two queues are not independent, but that the total times (waiting times plus service times) in each stage are independent.

Parallel Queues

A multichannel model has been defined previously as (1) possessing identical and independent servers in parallel and as (2) being further characterized by a channel selection rule which requires arrivals to select an empty channel without preference as long as such are available and to join the queue otherwise. Here we consider other parallel server arrangements. We will use "heterogeneous" to refer to those multichannel models in which the service time distributions are the same, but the mean service rates differ. Heterogeneous models may or may not possess the channel selection rule required of the standard ("homogeneous") multichannel models. Any multichannel model, whether homogeneous or heterogeneous, which does not possess this channel selection rule will be regarded as a network of queues in parallel.

Cohen [11], 1956, using the well-known result (e.g., [67]) that the sum of Poisson variates is again Poisson, has applied Burke's result to a fairly general network of M|M|c queues. In his model, a number of tandem queues, each with one or more M|M|c queues in series, are joined in a parallel array. The output of this array then becomes the input to a single channel (M|M|1) queue. Morris [60] presented several examples of M|M|c networks in his 1962 text.

Parallel, non-Poisson queues have been treated primarily within the framework of heterogeneous queues, priority arrivals, conditional channel selection, and other special models.

Daru [17], Gumbel [33], Gani [30], and Krishnamoorthi [47] have considered the case of heterogeneous M|M|c models. Daru suggested a number of methods for dealing with queues possessing servers of different efficiencies. Gumbel formulated a problem with heterogeneous servers and made an error study of the result when service rates are homogeneously approximated by an arithmetic average over all servers. Gani worked within the context of dam theory. He studied the first emptiness (first service completion) of two dams in parallel.

Krishnamoorthi's paper was both interesting and useful. He considered a heterogeneous M|M|2 system under two different queue disciplines. In the first, an arrival accepts service in the first free channel. If both channels are free, the arrival accepts service in channel one with probability π_1. In the second, up to m units will queue before channel one. The $(m+1)^{th}$ unit will take the first opportunity of entering channel two or moving into the m^{th} place. For both cases, Krishnamoorthi examined the equilibrium distribution of busy periods. As an application, he explained how one can choose the appropriate capacity of a relief channel if it is decided that an $M|M|1$ queue is overloaded.

Haight [34], 1958, initiated the study of two parallel queues with distinct waiting lines. Normally, an arrival chooses the shortest line; however, if the lines are of equal length, the arrival is assigned to the "near" queue. Once a line is chosen, the unit may not defect into the other line. Wilkins [91] noted an extension to a more general case where, if X is the length of the near queue, Y that of the other, and W(X,Y) is the probability of an arrival's joining the near one, we have

$$w(x,y) = \begin{cases} 1 , & x < y ; \\ w(x) , & x = y ; \\ 0 , & x > y . \end{cases}$$

Kingman [44] investigated the stability of this system and approximated the limiting, equilibrium joint probabilities of the two queue lengths.

Anker and Gafarian [3] have meticulously studied a multichannel model with Poisson input and heterogeneous negative exponential service time distributions. In their model, an arrival balks if the queue size is N and enters otherwise. He waits for service a maximum given time (random variable) and then reneges if his service has not yet been started. Steady-state results are obtained for most parameters of normal interest (except the output distribution).

Fagen and Riordan [22] studied the case of cooperating parallel channels with Poisson input, ordered service, and c identical servers which cooperate to service a unit. Once the fastest server has finished his portion of the service, the unit is ejected and another begins service. Two GI-type service distributions were considered, Erlangian service and a uniform distribution.

Romani [68] investigated a model with a variable number of channels. The queue is $M|M|1$, except whenever the queue size is N, a new arrival prompts the addition of a new server. All added servers are removed when the waiting line becomes empty. Phillips [63] studied a variation of Romani's model in which a maximum of c channels is allowed. A consequence of his model is that all queue sizes from 2 to c-1 are equiprobable. These models have important applications to production situations where some of the work is deferrable.

Queues with Feedback

Koenigsberg [45, 46] has modeled a coal-cutting problem as a set of tandem $M|M|1$ queues (operations) serving N units (mine faces) in rotation. Each operation has a negative exponential service time distribution and, after units one through N have been served, the sequence is repeated. This form of closed system is called "cyclic queues." Finch [26] has studied the same network. However, in his model, a unit returns to the j^{th} queue with probability p_j upon completing service at the last queue. Both authors obtained multivariate probability distributions for the number of units in each waiting line.

Benson and Gregory [6] have generalized Koenigsberg's model in another direction. In their network, arrivals from outside the system (also Poisson) may join the queue before any stage. Similarly, after service at a queue, a unit may go on to the next tandem stage or depart to a location outside the system.

Saaty [69; pp. 294-301] gave a systematic resumé of the work with M|M|1 cyclic queues and M|M|1 queues with feedback. He also discussed a single-server, cyclic, multiqueue model in which N queues are serviced (to emptiness) in rotation.

More recently, Takács [76] has investigated a non-Poisson feedback model. His model was a single queue of the type M|GI|1 in which departing units would immediately rejoin the queue with probability p. Distributions were obtained for queue size and a unit's total time in the system.

Other Queues

The moving single-server model of McMillan and Riordan [56] is not properly a network model. It is mentioned here because of the frequency with which it is encountered in production lines and other real situations. The model consists of a single server who processes randomly-spaced units on a moving assembly line. After the server finishes a unit, he moves back to the next unit in sequence without delay. The server's efficiency is measured by the length of time he remains before a given point on the line. That is, the production line has a barrier which absorbs the server if he crosses it. As an application, this barrier might be made to correspond to a missed service in a real production line. McMillan and Riordan obtained an expression for p(k,T), the probability that service is completed on k units before absorption when the server has started processing the first unit when it was T time units from the barrier. Karlin, Miller, and Prabhu [42] have pointed out that this model is equivalent to the single channel queue with Poisson input and service time distribution corresponding to the spacing distribution of McMillan and Riordan's model.

Interstage Flows

A number of papers on Poisson flows in networks appeared prior to Burke's result of 1956. These papers were primarily of the applied variety, but many did treat situations which might be encountered in general networks. We mention here only the works of Tanner and Boldyref because of their representative pertinence. Reference is made to Saaty [69], Takács [77], and Cox and Smith [16] for additional examples.

One problem of traditional interest to network analysts is determination of the expected maximum rate at which units could be processed through a complex network of operations. Boldyref's paper [7] investigated flow through a railroad network. Cars (in a switchyard) or trains have to be moved from one point in the network to another. Since individual tracks cannot be provided for each journey, queues will form at track junctions. Arrival rates and queue disciplines are, of course, dependent on the predetermined train schedules and priorities. For certain such railroad network examples under given scheduling conditions, Boldyref obtained the maximum steady-state flow through the network. Note that his results are pertinent to any network of queues where there is finite transportation capacity between stages.

Tanner [78] studied another problem in interstage flow, that of traffic interference. In Tanner's model, two freely flowing streams approach a single lane of finite length from opposite directions. It is assumed that units travel at constant speed and that starting and stopping times are negligible. Tanner obtained results for expected total delays and for expected interval lengths that one stream would control the lane. Note that the model is equally applicable to intersection traffic (where streams cross each other at, say, right angles) of the type found on city streets or plant aisles.

More recent work* has dealt with a number of other traffic flow problems. A particularly interesting example is Jewell's model [41] of traffic entries from a secondary road.

Another aspect of the flow merger problem involves the (parallel) arrival of independent streams at a single queue. If servicing is on a first-in-first-out basis, it is evident that the queue input is just the convolution of the independent number-of-arrivals-per-interval distributions. However, if the members of one stream have over-riding, preemptive priority of service, the problem becomes much more complex. White and Christe [80], Stephan [72, 73], Heathcote [35], and Ancker and Gafarian [2] have all studied this problem.

White and Christe and Stephan have obtained the steady-state equations of the dual input, preemptive priority $M|M|1$ model. The latter has also obtained the mean waiting time and other moments for the lower priority queue. Heathcote extended analysis to the time-dependent case. He determined the joint distribution of the numbers of priority and non-priority units in the system at time t for given initial conditions. Heathcote also investigated the busy period distribution for the non-priority units. Ancker and Gafarian have studied the superposition of independent Poisson streams with negative exponen-

---

*See Saaty [69], pp. 302-323 for a summary.

tial service times which depend upon the stream class to which a unit belongs.

CHAPTER III

OUTPUT OF MODELS WITH POISSON INPUT AND NEGATIVE EXPONENTIAL SERVICING

We will use $M_n|M_n|1$ to designate those steady-state queueing models which are characterized by a single state-dependent Poisson input density function (for the number of arrivals in time t) and a single service-station, or "channel," with a negative exponential service time density function. The state of a system is the total number of units in that system, that is, the sum of those in the queue and those being serviced. As we shall see, the assumption of a single channel is only tacit. All $M_n|M_n|1$ models are single channel in their mathematical form; however, multi-input and multichannel models do exist as a physical interpretation of certain $M_n|M_n|1$ models. This point will be elucidated in the discussion of the general $M_n|M_n|1$ model.

Preliminary Concepts

This section will present some of the equations and concepts which are basic to the analysis of $M_n|M_n|1$ queueing models. The results are well known; however, their development in the present context will clarify, shorten, and generally facilitate the discussion and analysis of this chapter.

Poisson Arrival Processes

The Poisson distribution is frequently used to represent input phenomena when arrivals occur essentially at "random" or when little or nothing is actually known about the input parameters. The basis for such usage stems from several causes: First, the Poisson distribution may be derived from very general assumptions (listed below) which agree quite well with our notion of what the probabilistic properties of a random phenomenon might be. Second, the ease with which computations may be effected overrides many possible objections.* Third, and most important, many real input systems (e.g. road traffic, telephone calls, restaurant customers, etc.) show excellent empirical correspondence to their Poisson analogues.

A number of well-known properties of the Poisson distribution will now be derived in the context of an arrival process.

Arrival Assumptions.

1. The probability of exactly one arrival in a small interval of length Δt is directly proportional to Δt. Denote this probability by $\lambda\Delta t$.

2. The probability of more than one arrival in a small interval of length Δt is of much higher order than Δt. Denote this probability by $\varepsilon_1(\Delta t)$.

3. $\lim\limits_{\Delta t\to 0} \frac{\lambda\Delta t}{\Delta t} = \lambda$; $\lim\limits_{\Delta t\to 0} \frac{\varepsilon_1(\Delta t)}{\Delta t} = 0$.

Derivation of Governing Equations. The event, a arrivals during the closed time interval $[0,t+\Delta t]$, may be thought of as occurring in one of three mutually exclusive and collectively exhaustive ways:

1. $\{a \text{ arrivals in } [0,t]\} \cap \{0 \text{ arrivals in } (t,t+\Delta t]\}$ (for each $a\geq 0$)

---

*E. C. Molina [57] has published extensive tables of individual and complementary cumulative terms of the Poisson distribution.

2. $\{a-1 \text{ arrivals in } [0,t]\} \cap \{1 \text{ arrival in } (t,t+\Delta t]\}$

(for each $a \geq 1$)

3. $\{a-x \text{ arrivals in } [0,t]\} \cap \{x \text{ arrivals in } (t,t+\Delta t]\}$

(for each $a \geq 2$; $2 \leq x \leq a$)

Letting $\Phi_a(t)$ denote the probability of exactly a arrivals during the interval $[0,t]$, the probabilities of these events are

1. $\Phi_a(t)[1-\lambda\Delta t-\varepsilon_1(\Delta t)]$ (each $a \geq 0$)
2. $\Phi_{a-1}(t)[\lambda\Delta t]$ (each $a \geq 1$)
3. $\Phi_{a-x}(t)[\varepsilon_x(\Delta t)]$ (each $a \geq 2$; $2 \leq x \leq a$)

where $\varepsilon_x(\Delta t)$ has been used to denote the probability of exactly x arrivals during a small interval of length Δt.

Applying the additive theorem for the probability of the union of independent events, we obtain

$$\left.\begin{aligned}
\Phi_o(t+\Delta t) &= \Phi_o(t)[1-\lambda\Delta t-\varepsilon_1(\Delta t)] \\
\Phi_1(t+\Delta t) &= \Phi_1(t)[1-\lambda\Delta t-\varepsilon_1(\Delta t)] + \Phi_o(t)[\lambda\Delta t] \\
\Phi_a(t+\Delta t) &= \Phi_a(t)[1-\lambda\Delta t-\varepsilon_1(\Delta t)] \\
&\quad + \Phi_{a-1}(t)[\lambda\Delta t] + \sum_{x=2}^{a} \Phi_{a-x}(t)\varepsilon_x(\Delta t)\ , \\
&\qquad a = 2,3,4,\cdots.
\end{aligned}\right\} \quad (1)$$

Transposing $\Phi_a(t)$ and dividing through by Δt results in

$$\left.\begin{aligned}
&\frac{\Phi_o(t+\Delta t) - \Phi_o(t)}{\Delta t} = -\lambda\Phi_o(t) - \Phi_o(t)\,\frac{\varepsilon_1(\Delta t)}{\Delta t} \\
&\frac{\Phi_1(t+\Delta t) - \Phi_1(t)}{\Delta t} = -\lambda\Phi_1(t) + \lambda\Phi_o(t) - \Phi_1(t)\,\frac{\varepsilon_1(\Delta t)}{\Delta t} \\
&\frac{\Phi_a(t+\Delta t) - \Phi_a(t)}{\Delta t} = -\lambda\Phi_a(t) + \lambda\Phi_{a-1}(t) - \Phi_a(t)\,\frac{\varepsilon_1(\Delta t)}{\Delta t} \\
&\qquad\qquad + \sum_{x=2}^{a} \Phi_{a-x}\,\frac{\varepsilon_x(\Delta t)}{\Delta t}\,, \qquad a = 2,3,4,\cdots.
\end{aligned}\right\} \quad (2)$$

Taking the limit as $\Delta t \to 0$ and noting that necessarily $\varepsilon_x(\Delta t) < \varepsilon_1(\Delta t)$, we obtain the equations governing the system:

$$\left.\begin{aligned}
&\frac{d\Phi_o(t)}{dt} = -\lambda\Phi_o(t) \\
&\frac{d\Phi_a(t)}{dt} = -\lambda\Phi_a(t) + \lambda\Phi_{a-1}(t), \qquad a = 1,2,3,\cdots.
\end{aligned}\right\} \quad (3)$$

Boundary conditions follow directly from the initial assumptions:

$$\Phi_o(0) = 1 \quad \text{and} \quad \Phi_a(0) = 0\ , \qquad a = 1,2,3,\cdots. \qquad (4)$$

<u>Number of Arrivals in Time t.</u> The set of differential equations (3), with boundary conditions (4), may be solved by a mathematical

induction to yield

$$\Phi_a(t) = \frac{(\lambda t)^a}{a!} e^{-\lambda t} , \qquad t \geq 0, \qquad a = 0,1,2,\cdots, \tag{5}$$

for the probability of a arrivals during the interval [0,t]. Equation (5) may be recognized as the Poisson distribution with mean λt. Its generating function is*

$$\Phi_z^g(t) \doteq \sum_{a=0}^{\infty} z^a \Phi_a(t) = e^{-\lambda t(1-z)} \tag{6}$$

Mean Arrival Rate. The mean arrival rate is the mean number of arrivals during an interval, divided by the length of that interval, i.e.

$$\text{Mean arrival rate} \doteq \frac{1}{T} E(a|T) = \frac{\lambda T}{T} = \lambda . \tag{7}$$

Time Between Arrivals. The probability density function, $g(t_{ba})$, for the time between arrivals, t_{ba}, may be obtained from consideration

---

*Recall that the generating function for a discrete probability density function is just its geometric, or "z-," transform. In general, we will use

$$f^g(z) \doteq \sum_{n=0}^{\infty} z^n f(n)$$

to denote the geometric transform of a discrete function, f(n). Reference is made to Beightler, *et al.* [4] for one of the few good lists of these transforms.

of Equation (5) with a = 0. Note that

$$\Phi_o(t) \doteq \Pr\left\{\begin{matrix}\text{exactly 0 arrivals}\\ \text{during } [0,t]\end{matrix}\right\} \equiv \Pr\left\{\begin{matrix}\text{no arrivals have occurred in time } t\\ \text{when observation is started at time } 0\end{matrix}\right\}.$$

Suppose that an (arbitrary) arrival has occurred at time 0. Then $\Phi_o(t)$ is just the probability that the waiting time until the next arrival exceeds t. Thus a complementary cumulative distribution function is

$$\Pr\{t_{ba} > t_o\} = e^{-\lambda t_o} , \tag{8}$$

for some $t_o \geq 0$. From this we obtain the cumulative distribution function,

$$\Pr\{0 \leq t_{ba} \leq t_o\} = 1 - e^{-\lambda t_o} , \tag{9}$$

and the density function,

$$\Pr\{t_{ba} = t_o\} = \frac{d}{dt_o}(1 - e^{-\lambda t_o}) = \lambda e^{-\lambda t_o} . \tag{10}$$

Hence, the desired function, $g(t_{ba})$, is given by

$$g(t_{ba}) = \lambda e^{-\lambda t_{ba}} , \quad t_{ba} \geq 0 . \tag{11}$$

Equation (11) may be recognized as a negative exponential distribution with mean $1/\lambda$. As expected, the mean interarrival interval is just the reciprocal of the mean arrival rate, λ. The generating function, $g^e(s)$, for $g(t_{ba})$ is*

$$g^e(s) \doteq \int_0^\infty g(t_{ba})e^{-st_{ba}}\,dt_{ba} = \frac{\lambda}{s+\lambda}\ . \tag{12}$$

Negative Exponential Servicing

The assumptions leading to a negative exponential distribution for the length of service time on a unit are essentially equivalent to those of the Poisson input process. The one important difference in the present situation is the provision that there are units available for servicing at all times. The derivation proceeds as follows:

Service Assumptions.

1. Given continuous servicing, the probability of exactly one service completion in a small interval of length Δt is directly proportional to Δt. Denote this probability by $\mu\Delta t$.

2. Given continuous servicing, the probability of more than one service completion in a small interval of length Δt is of much higher order than Δt. Denote this probability by $\varepsilon_2(\Delta t)$.

---

*Recall that the generating function of a continuous probability density function is its unilateral Laplace, or negative exponential, transform. In general, we will use

$$f^e(s) \doteq \int_0^\infty f(x)e^{-sx}dx$$

to denote the Laplace transform of a continuous function, $f(x)$.

3. $\lim_{\Delta t \to 0} \frac{\mu \Delta t}{\Delta t} = \mu$; $\lim_{\Delta t \to 0} \frac{\varepsilon_2(\Delta t)}{\Delta t} = 0$.

Derivation of Governing Equations. Define $\Psi_b(t)$ as the conditional probability of exactly b service completions during the interval [0,t], given continuous servicing throughout the interval [0,t]. Then, replacing λ by μ, $\varepsilon_1(\Delta t)$ by $\varepsilon_2(\Delta t)$, and $\Phi_a(t)$ by $\Psi_b(t)$ in the derivation of the Poisson input process, we obtain

$$\Psi_b(t) = \frac{(\mu t)^b}{b!} e^{-\mu t} , \quad t \geq 0, \quad b = 0,1,2,\cdots. \tag{13}$$

The mean of this Poisson distribution is μt and its generating function is

$$\Psi_z^g(t) \doteq \sum_{b=0}^{\infty} z^b \Psi_b(t) = e^{-\mu t(1-z)} . \tag{14}$$

Mean Service Rate. The mean service rate is defined as the conditional expectation of the number of services completed in one time unit, given that servicing is going on throughout the entire time, i.e.

$$\text{Mean service rate} \doteq \frac{1}{T} E(b | T, \text{ continuous servicing}) = \frac{\mu T}{T} = \mu. \tag{15}$$

Service Times. $\Psi_o(t)$ may be interpreted as the probability that service time on a unit exceeds t when observation starts at time 0. Then, in analogy to Equations (8-10), we obtain for service time, t_s,

$$\Pr\{t_s > t_o\} = e^{-\mu t_o} , \tag{16}$$

$$\Pr\{0 \le t_s \le t_o\} = 1 - e^{-\mu t_o} , \quad \text{and} \tag{17}$$

$$\Pr\{t_s = t_o\} = \frac{d}{dt_o}\,(1 - e^{-\mu t_o}) = \mu e^{-\mu t_o} . \tag{18}$$

Thus, the probability density function, $v(t_s)$ for the time, t_s, required to service a single (arbitrary) unit is

$$v(t_s) = \mu e^{-\mu t_s} , \; t_s \ge 0 . \tag{19}$$

Equation (19) is just the negative exponential distribution with mean, μt_s, and generating function,

$$v^e(s) = \frac{\mu}{s+\mu} . \tag{20}$$

The negative exponential distribution is frequently used to model human service systems and (primarily for ease) other stochastic service systems. The representation may seem reasonable enough when regarded in reference to the underlying assumptions and number-of-completed-services probabilities; however, an implication of Equation (19) is that a unit can be serviced in zero time with finite probability μ. This drawback may not seriously impair the usefulness of the representation, especially in those cases where primary interest is focused on the state of (number of units in) a queueing system. In many real queueing systems, it is impractical to measure the input and/or output parameters. While a Poisson input and negative exponential service times may not separately be good representations of their real analogues, taken together they may result in excellent results for the state parameters of the system. This has certainly been the case in a number of empirical studies.

Another reason for using a negative exponential service time distribution in conjunction with a Poisson input is the resulting computational and algebraic ease. This will be demonstrated below in the derivation of a general $M_n|M_n|1$ queueing model. First, however, we will demonstrate the important Markovian property of this distribution.

Markovian Property of the Negative Exponential Distribution

In general, we say that a process is Markovian in nature if the probability of some future event is dependent, at most, upon the present state of the system. That is, a Markovian process has no "memory" of its previous behavior. The fact that the negative exponential distribution has this property will be extremely important to waiting time determination in our analysis of the output of $M_n|M_n|1$ queueing systems.

Consider the negative exponential distribution,

$$f(t) = ke^{-kt} , \; t \geq 0 . \tag{21}$$

Suppose conditions are such that f(t) may be interpreted as the probability that an arbitrary inter-event interval, started at time 0, will terminate at time t. Further, suppose it is known that the interval has not terminated by time t_o. With such partial information, we obtain

$$f(t|t>t_o) \doteq \frac{f(t)}{1 - \int_o^{t_o} f(t)dt} = ke^{-k(t-t_o)} , \quad t \geq t_o, \qquad (22)$$

as the conditional probability that the interval terminates at time t. If the origin is moved to t_o and $t - t_o$ is set equal to t', then Equation (22) becomes

$$f(t') = ke^{-kt'} , \quad t' \geq 0 . \qquad (23)$$

The implications of this result are important. In particular, note that $g(t_{ba})$ in Equation (11) may now be interpreted as the probability density function for waiting time until the next arrival when observation is started at some t = 0 anywhere within an interarrival interval. A similar interpretation holds for Equation (19).

<u>The General $M_n|M_n|1$ Model</u>

<u>Characteristics of the Model</u>.

1. Poisson input--with mean arrival rate λ_n when there are n units in the system.

2. Negative exponential servicing--with mean rate μ_n when there are n units in the system.

3. Steady-state conditions.

<u>Amended Poisson Assumptions</u>. Assumptions for a Poisson arrival process and for a negative exponential service time distribution were

given separately when these processes were discussed above. Taken together, these six form the assumptions for the $M_n|M_n|1$ model if, in each case, the given assumption is prefaced by "Given n units in the system" and λ_n, $\varepsilon_{1,n}(\Delta t)$, μ_n, and $\varepsilon_{2,n}(\Delta t)$ are written for λ, $\varepsilon_1(\Delta t)$, μ, and $\varepsilon_2(\Delta t)$, respectively.

These six assumptions are entirely sufficient for mathematical development of the general $M_n|M_n|1$ model. It should be noted that, while the model is tacitly assumed to possess only a single input source and a single service station (channel), it is possible to obtain multichannel models by suitable specification and interpretation of the servicing parameters, μ_n.* For example, one set of multichannel models (designated $M_n|M|c$ is obtained by letting μ_n be proportional to n over some finite range of n, say $\mu_n = n\mu$ for $1 \leq n \leq c$, and setting $\mu_n = c\mu$ for $n > c$. The resulting model may be interpreted as having c identical and independent service stations in parallel, each possessing a negative exponential service time distribution with mean rate μ; a single, state-dependent Poisson input; and a channel-selection rule requiring new arrivals to enter one of the empty stations, without preference (equiprobably), as long as they are available, and to join the queue otherwise.

Derivation of Governing Equations. The event, n units in the system at time $t + \Delta t$, can be written as the union of a number of

---

*It is quite evident that a more general multi-channel could be constructed by simply specifying the number of channels, their service time distributions (negative exponential, if desired), and a channel-selection rule to distribute their common input (Poisson, if desired). However, in this case, it proves more convenient to speak of the resulting model as a network of queueing models in parallel.

mutually exclusive and collectively exhaustive subevents, i.e.

$$
\begin{aligned}
\{S_n(t+\Delta t)\} = {} & [\{S_n(t)\} \cap \{a_o(\Delta t)\} \cap \{d_o(\Delta t)\}] \\
& \bigcup [\{S_{n-1}(t)\} \cap \{a_1(\Delta t)\} \cap \{d_o(\Delta t)\}] \\
& \bigcup [\{S_{n+1}(t)\} \cap \{a_o(\Delta t)\} \cap \{d_1(\Delta t)\}] \\
& \bigcup_{x=2}^{n} [\{S_{n-x}(t)\} \cap \{a_x(\Delta t)\} \cap \{d_o(\Delta t)\}] \\
& \bigcup_{x=2}^{\infty} [\{S_{n+x}(t)\} \cap \{a_o(\Delta t)\} \cap \{d_x(\Delta t)\}] \\
& \bigcup_{k=1}^{n} \bigcup_{x=1}^{\infty} [\{S_{n-k}(t)\} \cap \{a_{k+x}(\Delta t)\} \cap \{d_x(\Delta t)\}] \\
& \bigcup_{k=0}^{\infty} \bigcup_{x=1}^{\infty} [\{S_{n+k}(t)\} \cap \{a_x(\Delta t)\} \cap \{d_{k+x}(\Delta t)\}] \ ,
\end{aligned}
\tag{24}
$$

where $S_n(t)$ is the statement "n units are in the system at time t," $a_x(\Delta t)$ is the statement "x arrivals occur during an interval of length Δt," and $d_x(\Delta t)$ is the statement "x departures (service completions) occur during an interval of length Δt."

Let $P_n(t)$ denote the probability of n units in the system at time t. Then the probability relationships implied by Equation (24) are

$$\left.\begin{aligned}
P_o(t+\Delta t) &= P_o(t)[(1-\lambda_o\Delta t)(1)] + P_1(t)[(1-\lambda_1\Delta t)(\mu_1\Delta t)] + \\
&\qquad \text{(second and higher order terms in } \Delta t\text{)}, \ n = 0, \quad \text{and} \\
P_n(t+\Delta t) &= P_n(t)[(1-\lambda_n\Delta t)(1-\mu_n\Delta t)] + P_{n-1}(t)[(\lambda_{n-1}\Delta t)(1-\mu_{n-1}\Delta t)] \\
&\qquad + P_{n+1}(t)[(1-\lambda_{n+1}\Delta t)(\mu_{n+1}\Delta t)] + \text{(second and higher} \\
&\qquad \text{order terms in } \Delta t\text{)}, \quad n = 1,2,3,\cdots.
\end{aligned}\right\} \quad (25)$$

Transposing $P_n(t)$ and dividing by Δt, and taking the limit as $\Delta t \to 0$, we obtain the dynamic equations of the system,

$$\left.\begin{aligned}
\frac{dP_o(t)}{dt} &= -\lambda_o P_o(t) + \mu_1 P_1(t), \quad n = 0, \text{ and} \\
\frac{dP_n(t)}{dt} &= -(\lambda_n+\mu_n)P_n(t) + \lambda_{n-1}P_{n-1}(t) + \mu_{n+1} P_{n+1}(t), \\
&\qquad n = 1,2,3,\cdots.
\end{aligned}\right\} \quad (26)$$

If the system started in state k at t = 0, the initial conditions are

$$P_k(0) = 1 \qquad \text{and} \qquad P_n(0) = 0, \quad n \neq k . \qquad (27)$$

Solutions to Equations (26) are not easily obtained. Indeed, the question of existence and uniqueness of the $P_n(t)$ is far from trivial. The primary difficulty lies in the fact that this infinite set of differential equations must be solved simultaneously rather than recursively as was the case with Equations (3) or (13). However, for most practical models (including all those to be presented in this chapter), it may safely be assumed that unique, regular solutions to Equations (26) do exist.*

Steady-State Solution. Express $P_n(t)$ in the general form,

$$P_n(t) = p_n + f_n(t) \ , \tag{28}$$

where p_n and $f_n(t)$ are, respectively, the time invarient and time varient parts of $P_n(t)$. It can be shown [23] that $f_n(t)$ is a purely transient term, that is,

$$\lim_{t\to\infty} \frac{dP_n(t)}{dt} = 0 \tag{29}$$

and

$$\lim_{t\to\infty} p_n(t) = p_n \ , \tag{30}$$

independent of the initial conditions. It is important to note that Equation (30) does not imply a cessation of system activity as $t\to\infty$. Rather, the implication is only that the average proportion of time spent in each state under steady-stage conditions is independent of the starting state.

---

*See Feller [24], pp. 407-411, for a more detailed discussion. A short annotated bibliography of papers dealing with non-regular solutions is given in a footnote.

Taking the limit as $t \to \infty$ in Equations (26) results in

$$\left.\begin{aligned} 0 &= -\lambda_o p_o + \mu_1 p_1 \\ \text{and} \qquad 0 &= -(\lambda_n+\mu_n)p_n + \lambda_{n-1}p_{n-1} + \mu_{n+1}p_{n+1}, \qquad n = 1,2,3,\cdots \end{aligned}\right\} \quad (31)$$

Equations (31) can be solved by recurrence relationships and induction to yield

$$p_n = \frac{\lambda_0\lambda_1\lambda_2\cdots\lambda_{n-1}}{\mu_1\mu_2\mu_3\cdots\mu_n} p_o, \quad n = 1,2,3,\cdots. \qquad (32a)$$

if all of the μ_n are non-zero for $n \geq 1$. If some μ_i is zero, then the solution is

$$p_n = \begin{cases} 0 & , \quad n = 0,1,2,\cdots,i-1 ; \\ \dfrac{\lambda_i\lambda_{i+1}\cdots\lambda_{n-1}}{\mu_{i+1}\mu_{i+2}\cdots\mu_n} & , \quad n = i,i+1,i+2,\cdots. \end{cases} \qquad (32b)$$

The case of a $\lambda_i = 0$ is handled implicitly in both forms of Equation (32). That is, $p_n = 0$ for $n > i$.

Since μ_n is rarely zero for $n \geq 1$ in a real situation, we need only treat the cases where all μ_n, for $n \geq 1$, are non-zero. However, should one desire an output distribution for some model where the μ_n are not all non-zero, then he can obtain it by starting with the initial, tabular form of the output distribution (see Tables 1-3) and applying

Equation (32b), rather than Equation (32a), throughout the given derivation. All results which follow in this chapter are based on the assumption that Equation (32a) holds. By the certainty theorem for a regular solution,

$$1 = \sum_{n=0}^{\infty} p_n = p_o \sum_{n=0}^{\infty} A_n , \tag{33}$$

where

$$A_n = \begin{cases} 1 & , \quad n = 0 ; \\ \prod_{r=1}^{n} \frac{\lambda_{r-1}}{\mu_r} & , \quad n = 1,2,3,\cdots . \end{cases} \tag{34}$$

Hence,

$$p_o = [\sum_{n=0}^{\infty} A_n]^{-1} . \tag{35}$$

Note that Equation (35) holds only for finite and non-zero p_o. If $p_o = 0$, then $p_n = p_o = 0$ for all n. This latter case corresponds to a discrete rectangular probability distribution over the positive integers of zero height, infinite width, and unitary area. For such equiprobable states, the notion of a steady-state model breaks down since the state of the system would be randomly distributed. In the models which follow, we will consider only those cases for which p_o is non-zero.

We are now in a position to ascertain the output distributions of several $M_n|M_n|1$ models. In each case, it will prove convenient to work in terms of the departure epochs and interdeparture intervals. This form of development and result is standard in the literature; however, should the reader desire a probabilistic expression for the number of departures in time t, he can obtain it by reversing the argument of Equations (8-11) or through other well-known procedures.

Output of the M|M|c Model

The physical analogue of the M|M|c model has previously been described as a set of c service stations in parallel, possessing identical, but independent, negative exponential service time distributions. Arrivals are from a single, state-independent Poisson input and must choose an empty station without preference among empty stations as long as such are available.

The mathematical model is completely specified by

$$\left.\begin{array}{lll} n = 0 & \lambda_o = \lambda & \mu_o = 0 \\ n = 1,2,\cdots,c-1 & \lambda_n = \lambda & \mu_n = n\mu \\ n = c,c+1,\cdots & \lambda_n = \lambda & \mu_n = c\mu \end{array}\right\} \tag{36}$$

The steady-state probabilities may be obtained from Equation (32a). These are

$$P_n = \begin{cases} \dfrac{1}{n!}\left(\dfrac{\lambda}{\mu}\right)^n P_o & , \quad n = 1,2,\cdots,c-1 \; ; \\ \dfrac{1}{c!\; c^{n-c}}\left(\dfrac{\lambda}{\mu}\right)^n P_o & , \quad n = c,c+1,\cdots . \end{cases} \tag{37}$$

From Equation (35), we obtain

$$p_o = [\sum_{n=0}^{c-1} \frac{1}{n!} (\frac{\lambda}{\mu})^n + \sum_{n=c}^{\infty} \frac{1}{c!\ c^{n-c}} (\frac{\lambda}{\mu})^n]^{-1} . \qquad (38)$$

Since we are interested only in the non-saturated case ($p_\infty \neq 1$), we set the condition,

$$\frac{\lambda}{c\mu} < 1 , \qquad (39)$$

and evaluate the second summation in Equation (38) to yield

$$p_o = [\sum_{n=0}^{c-1} \frac{1}{n!} (\frac{\lambda}{\mu})^n + \frac{(\frac{\lambda}{\mu})^c}{c!\ (1- \frac{\lambda}{c\mu})}]^{-1} . \qquad (40)$$

Rather than proceeding directly to the derivation of an output distribution for the general c-channel case, we will first demonstrate a combinatorial rational, one used extensively throughout this chapter, on the more simple $M|M|1$ and $M|M|2$ models. We will retain our earlier notation; however, $v(t_s|n\mu)$ and $\Psi_b(t_s|n\mu)$ will replace $v(t_s)$ and $\Psi_b(t_s)$, respectively, in acknowledgment of the state-dependency of the service rate. The symbol t_{bd} will be used to denote the time between departures and $h(t_{bd})$, its density function.

<u>Output of $M|M|1$</u>

Suppose a departure has just occurred at time $t = 0$. The system is now empty (n=0) with probability, p_o, <u>or</u> the system now has one or more units in it ($n \geq 1$) with probability, $(1-p_o)$. In the latter case, the next departure will occur as soon as the next unit in line can be serviced. Thus,

$$t_{bd} = t_s \qquad \text{with probability } (1-p_o) \tag{41}$$

and

$$h(t_{bd}) = v(t_{bd}|\mu) = \mu e^{-\mu t_{bd}} \tag{42}$$

$$\text{with probability } (1-p_o) \; .$$

In the former, the next departure will occur as soon as the next arrival can be serviced. From the results of Equation (23) and associated remarks, we know that the waiting time for an arrival has the same distribution as the time between arrivals. Hence,

$$t_{bd} = t_{ba} + t_s \qquad \text{with probability} \quad p_o \; . \tag{43}$$

Thus $h(t_{bd})$ is just the convolution of $g(t_{ba})$ and $v(t_s|\mu)$, viz.,

$$h(t_{bd}) = g(t_{bd}) * v(t_{bd}|\mu) = \int_o^{t_{bd}} g(x)v(t_{bd}-x|\mu)dx \tag{44}$$

$$\text{with probability } p_o \; .$$

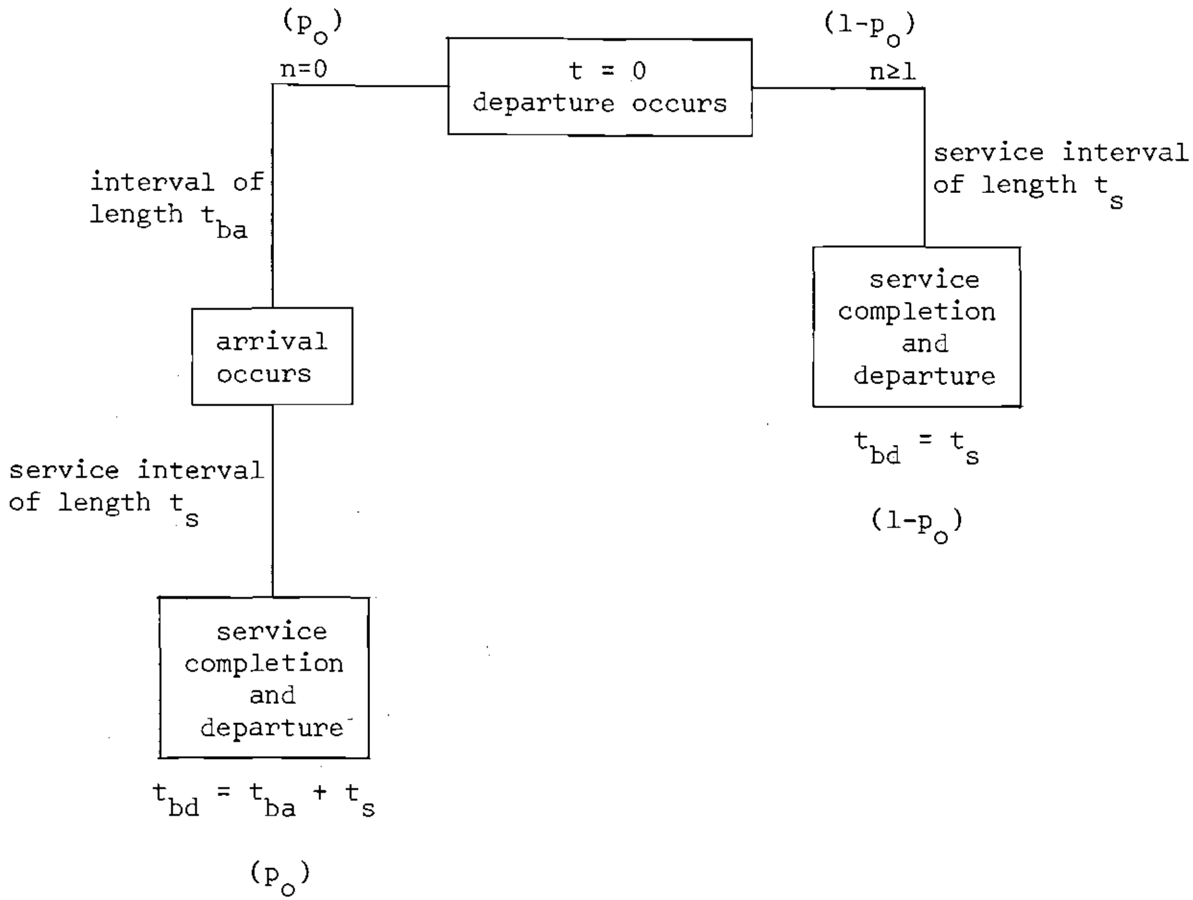


Figure 5. Interdeparture Events for M|M|1.

These interdeparture events are summarized in Figure 5.

Since the two sets of events are independent and collectively exhaustive of all possible interdeparture behavior, we may write for their union,

$$h(t_{bd}) = (1-p_o)v(t_{bd}|\mu) + (p_o)[g(t_{bd})*v(t_{bd}|\mu)] \; . \tag{45}$$

Taking the Laplace transform of Equation (45), we obtain

$$\begin{aligned} h^e(s) &= (1-p_o)v^e(s|\mu) + (p_o)g^e(s)v^e(s|\mu) \\ &= (\frac{\lambda}{\mu})\,\frac{\mu}{s+\mu} + (1-\frac{\lambda}{\mu})(\frac{\lambda}{s+\lambda})(\frac{\mu}{s+\mu}) \\ &= \frac{\lambda}{s+\lambda}\,[\frac{s+\lambda}{s+\mu} + \frac{(\mu-\lambda)}{s+\mu}\,] \\ &= \frac{\lambda}{s+\lambda} \; . \end{aligned} \tag{46}$$

The inverse transform of Equation (46) is

$$h(t_{bd}) = \lambda e^{-\lambda t_{bd}} \; , \qquad t_{bd} \geq 0 \; , \tag{47}$$

which can be seen to be identical with the input distribution. We will soon see that Equation (47) holds for the multichannel case as well.

Output of M|M|2

The derivation of an output distribution for M|M|2 follows the pattern of the $M|M|1$ derivation. We suppose that a departure has just occurred at time $t = 0$, leaving the system with zero, one, or two or more units. We trace five mutually exclusive and collectively exhaustive sets of events to the occurrence of the next departure as shown in Figure 6. The essential difference between this derivation and that for $M|M|1$ is the close attention which must be paid to changing service rates and the occurrence of arrivals in order to ensure the mutual independence of the chosen events.

Several things should be noted about Figure 6. First, the service time has been given an extra subscript to indicate the service rate in cases where it was not inherently specified. As before, the Markovian property of the negative exponential distribution has been used to express the waiting time for an arrival as t_{ba}. In addition, this property has been used to obtain the service time distribution at rate 2μ when a second unit enters the queue while service is in progress on a unit. Finally, note that certain branches were weighted by the probability of no arrivals (or no departures) to ensure independence. On other branches, there was no such weighting since additional arrivals (with $n \geq 2$) could not affect the service rate.

Formulation of the output distribution is shown in Table 1.

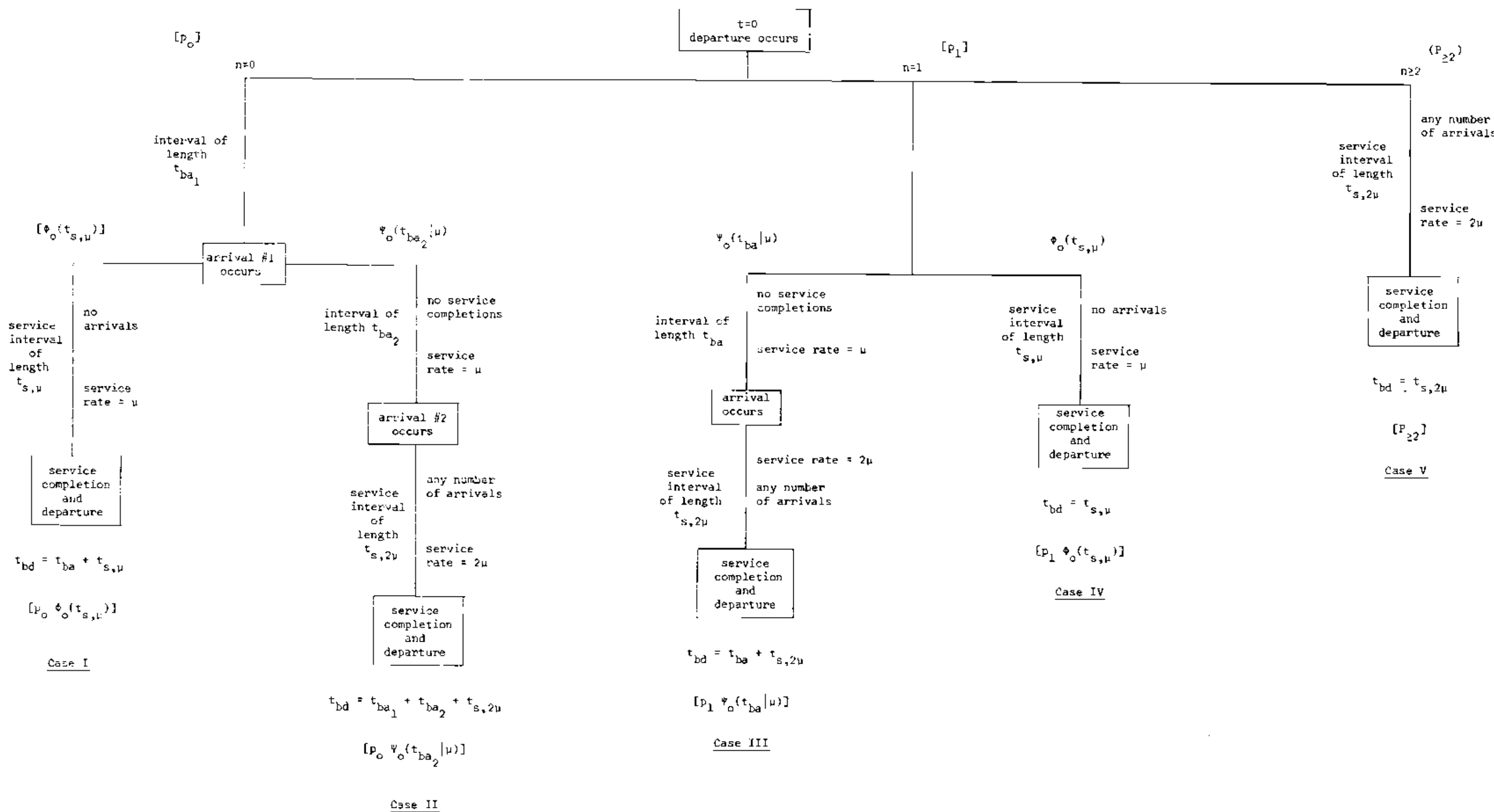


Figure 6. Interdeparture Events for M|M|2.

Table 1. Formulation of Output Distribution for $M|M|2$.

n	t_{bd}	With Probability	$p_n h(t_{bd})$
0	$t_{ba}+t_{s,\mu}$	$p_o \Phi_o(t_{s,\mu})$	$p_o\ g(t_{bd})*[\Phi_o(t_{bd})v(t_{bd}\|\mu)]$
0	$t_{ba_1}+t_{ba_2}+t_{s,2\mu}$	$p_o \Psi_o(t_{ba_2}\|\mu)$	$p_o\ g(t_{bd})*[\Psi_o(t_{bd}\|\mu)g(t_{bd})]*v(t_{bd}\|2\mu)$
1	$t_{s,\mu}$	$p_1 \Phi_o(t_{s,\mu})$	$p_1 \Phi_o(t_{bd})v(t_{bd}\|\mu)$
1	$t_{ba}+t_{s,2\mu}$	$p_1 \Psi_o(t_{ba}\|\mu)$	$p_1[\Psi_o(t_{bd}\|\mu)g(t_{bd})]*v(t_{bd}\|2\mu)$
≥2	$t_{s,2\mu}$	$P_{\geq 2}$	$P_{\geq 2}\ v(t_{bd}\|2\mu)$

$h(t_{bd})$ for the union of the five events is just the sum of the terms in the last column.

$$h(t_{bd}) = p_o\ g(t_{bd})*[\Phi_o(t_{bd})v(t_{bd}|\mu)] \tag{48}$$

$$+ p_o\ g(t_{bd})*[\Psi_o(t_{bd}|\mu)g(t_{bd})]*v(t_{bd}|2\mu)$$

$$+ p_1\ \Phi_o(t_{bd})v(t_{bd}|\mu) + p_1[\Psi_o(t_{bd}|\mu)g(t_{bd})]*v(t_{bd}|2\mu)$$

$$+ P_{\geq 2} v(t_{bd}|2\mu)$$

$$= p_o \int_o^{t_{bd}} g(t_{bd}-x)\ \Phi_o(x) v(x|\mu) dx$$

$$+ p_o \int_o^{t_{bd}} \int_o^{t_{bd-x}} g(t_{bd}-x-y) \Psi_o(x|\mu) g(x) v(y|2\mu) dydx$$

$$+ p_1\ \Phi_o(t_{bd}) v(t_{bd}|\mu) + p_1 \int_o^{t_{bd}} \Psi_o(t_{bd}-x|\mu) g(t_{bd}-x) v(x|2\mu) dx$$

$$+ P_{\geq 2} v(t_{bd}|2\mu)$$

$$= p_o \int_o^{t_{bd}} \lambda e^{-\lambda(t_{bd}-x)}\ e^{-\lambda x}\ \mu e^{-\mu x}\ dx$$

$$+ p_o \int_o^{t_{bd}} \int_o^{t_{bd}-x} \lambda e^{-\lambda(t_{bd}-x-y)} e^{-\mu x}\ \lambda e^{-\lambda x}\ 2\mu e^{-2\mu y}\ dydx$$

$$+ p_1 e^{-\lambda t_{bd}}\ \mu e^{-\mu t_{bd}} + p_1 \int_o^{t_{bd}} e^{-\mu(t_{bd}-x)}\ \lambda e^{-\lambda(t_{bd}-x)}\ 2\mu e^{-2\mu x}\ dx$$

$$+ P_{\geq 2}\ 2\mu e^{-2\mu t_{bd}}\ .$$

Its Laplace transform is

$$h^e(s) = p_o\left(\frac{\lambda}{s+\lambda}\right)\left(\frac{\mu}{s+\lambda+\mu}\right) + p_o\left(\frac{\lambda}{s+\lambda}\right)\left(\frac{\lambda}{s+\lambda+\mu}\right)\left(\frac{2\mu}{s+2\mu}\right) \qquad (49)$$

$$+ p_1\left(\frac{\mu}{s+\lambda+\mu}\right) + p_1\left(\frac{\lambda}{x+\lambda+\mu}\right)\left(\frac{2\mu}{x+2\mu}\right)$$

$$+ P_{\geq 2} \left(\frac{2\mu}{s+2\mu}\right) .$$

Substituting from Equations (37-38) with c = 2,

$$h^e(s) = p_o\left[\frac{\lambda\mu}{(s+\lambda)(s+\lambda+\mu)} + \frac{2\mu\lambda^2}{(s+\lambda)(s+\lambda+\mu)(s+2\mu)} + \frac{\lambda}{(s+\lambda+\mu)}\right. \tag{50}$$

$$\left. + \frac{2\lambda^2}{(s+\lambda+\mu)(s+2\mu)} + \frac{2\lambda^2}{(2\mu-\lambda)(s+2\mu)}\right]$$

$$= \left(\frac{\lambda}{s+\lambda}\right)\left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{\mu(2\mu-\lambda)}\right)^{-1} \left[\frac{\mu}{s+\lambda+\mu} + \frac{2\lambda\mu}{(s+\lambda+\mu)(s+2\mu)}\right.$$

$$\left. + \frac{(s+\lambda)}{(s+\lambda+\mu)} + \frac{2\lambda(s+\lambda)}{(s+\lambda+\mu)(s+2\mu)} + \frac{2\lambda(s+\lambda)}{(2\mu-\lambda)(s+2\mu)}\right]$$

$$= \left(\frac{\lambda}{s+\lambda}\right)\left(\frac{2\mu+\lambda}{2\mu-\lambda}\right)^{-1}\left[1 + \frac{2\lambda}{s+2\mu} + \frac{2\lambda(s+\lambda)}{(2\mu-\lambda)(s+2\mu)}\right] = \frac{\lambda}{s+\lambda} .$$

Hence,

$$h(t_{bd}) = \lambda e^{-\lambda t_{bd}}, \quad t_{bd} \geq 0 . \tag{51}$$

Again we obtain the Poisson result for the output.

<u>Output of M|M|c</u>

The formulation proceeds in the same way as for M|M|1 and M|M|2; however, in this case, there are [(1/2)(c+1)(c+2)-1]* mutually exclusive and collectively exhaustive interdeparture events to treat. The results

---

*This total is easily determined by direct count of the terms in Figure 7.

of this formulation are summarized in Table 2.

As before, $h^e(s)$ for the union of the events is just the sum of the entries in the last column. At this point, it becomes convenient to adopt an abbreviated notation. Let:

$$x_k = s + \lambda + k\mu \ , \qquad k = 0,1,\cdots,c-1 \ ; \tag{52}$$

$$x_c = s + c\mu = x_o + c\mu - \lambda \ ; \tag{53}$$

$$\pi_{c-1} = (s+\lambda)(s+\lambda+\mu)\cdots[s+\lambda+(c-1)\mu] = x_o x_1 \cdots x_{c-1} \ ; \tag{54}$$

and

$$\pi_c = (s+c\mu)\pi_{c-1} = x_o x_1 \cdots x_{c-1} x_c \ . \tag{55}$$

Using this notation, $h^e(s)$ becomes

$$\begin{aligned} h^e(s) = \frac{1}{\pi_c} \{ & p_o[0 + \lambda\mu(x_2\cdots x_c) + 2\mu\lambda^2(x_3\cdots x_c) + \cdots + (c-1)\mu\lambda^{c-1}x_c + c\mu\lambda^c] \\ & + p_1[\mu(x_o x_2\cdots x_c) + 2\mu\lambda(x_o x_3\cdots x_c) + \cdots + (c-1)\mu\lambda^{c-2}x_o x_c + c\mu\lambda^{c-1}x_o] \\ & + p_2[2\mu(x_0 x_1 x_3\cdots x_c) + 3\mu\lambda(x_0 x_1 x_4\cdots x_c) + \cdots + (c-1)\mu\lambda^{c-3}x_0 x_1 x_c \\ & \quad + c\mu\lambda^{c-2}x_o x_1] + \cdots \\ & + p_{c-2}[(c-2)\mu(x_o\cdots x_{c-3}x_{c-1}x_c) + (c-1)\mu\lambda(x_o\cdots x_{c-3}x_c) \\ & \quad + c\mu\lambda^2(x_o\cdots x_{c-3})] \end{aligned} \tag{56}$$

$$+ p_{c-1}[(c-1)\mu(x_o \cdots x_{c-2}x_c) + c\mu\lambda(x_o \cdots x_{c-2})]$$

$$+ P_{\geq c}[c\mu(x_o \cdots x_{c-1})]\} ,$$

where the 0 has been added as the first term in the coefficient of (p_o/π_c) for reasons that will shortly become clear.

Now, let $q(n,c-j)$ denote the $(j+1)^{th}$ term in the coefficient of (p_n/π_c) in the above expression $(n=0,1,\cdots c,c-1;\ j=0,1,\cdots,c-n)$. Let $q(c,c)$ be the coefficient of $(P_{\geq c}/\pi_c)$. It can be seen that with the exception of $q(0,c)$, $q(n,c-j)$ is just that term in the coefficient of (p_n/π_c) which has s^{c-j} as its highest power of s. Note also that the identifiers, $(n,c-j)$, correspond to the event entries in Table 2 with which the $q(n,c-j)$ are associated.

Arrange the terms of $h^e(s)$ in a triangular matrix array as shown in Figure 7. Each term should be listed only once.

With the aid of Equations (37; 52-55) sum the terms on the first diagonal (d_1):

$$d_1 = \frac{1}{\pi_c}(p_o q(0,c-1) + p_1 q(1,c)) \tag{57}$$

$$= \frac{1}{\pi_c}(p_o \lambda\mu x_2 \cdots x_c + p_1 \mu x_o x_2 \cdots x_c)$$

$$= \frac{x_2 \cdots x_c}{\pi_c}(\lambda\mu p_o + \lambda p_o x_o)$$

$$= \frac{p_o \lambda \pi_{c-1}}{\pi_c} = p_o\left(\frac{\lambda}{s+\lambda}\right) .$$

Table 2. Formulation of Output Distribution for M|M|c

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
0,1	0	$t_{ba}+t_{s,\mu}$	$p_o\Phi_o(t_{s,\mu})$	$p_o(\frac{\lambda}{s+\lambda})(\frac{\mu}{s+\mu+\lambda})$
0,2	0	$t_{ba_1}+t_{ba_2}+t_{s,2\mu}$	$p_o\Psi_o(t_{ba_2}\|\mu)\Phi_o(t_{s,2\mu})$	$p_o(\frac{\lambda}{s+\lambda})(\frac{\lambda}{s+\lambda+\mu})(\frac{2\mu}{s+2\mu+\lambda})$
0,3	0	$t_{ba_1}+t_{ba_2}+t_{ba_3}+t_{s,3\mu}$	$p_o\Psi_o(t_{ba_2}\|\mu)\Psi_o(t_{ba_3}\|2\mu)\Phi_o(t_{s,3\mu})$	$p_o(\frac{\lambda}{s+\lambda})(\frac{\lambda}{s+\lambda+\mu})(\frac{\lambda}{s+\lambda+2\mu})(\frac{3\mu}{s+3\mu+\lambda})$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
0,(c-1)	0	$(\sum_{x=1}^{c-1} t_{ba_x})+t_{s,(c-1)\mu}$	$p_o[\prod_{x=2}^{c-1}\Psi_o(t_{ba_x}\|(x-1)\mu)]\Phi_o(t_{s,(c-1)\mu})$	$p_o(\frac{\lambda}{s+\lambda})\cdots(\frac{\lambda}{s+\lambda+(c-2)\mu})(\frac{(c-1)\mu}{s+(c-1)\mu+\lambda})$
0,c	0	$(\sum_{x=1}^{c} t_{ba_x})+t_{s,c\mu}$	$p_o[\prod_{x=2}^{c}\Psi_o(t_{ba_x}\|(x-1)\mu)]$	$p_o(\frac{\lambda}{s+\lambda})\cdots(\frac{\lambda}{s+\lambda+(c-1)\mu})(\frac{c\mu}{s+c\mu})$
1,0	1	$t_{s,\mu}$	$p_1\Phi_o(t_{s,\mu})$	$p_1(\frac{\lambda}{s+\mu+\lambda})$

Table 2. Formulation of Output Distribution for M|M|c (Continued)

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
1,1	1	$t_{ba}+t_{s,2\mu}$	$p_1\Psi_o(t_{ba}\|\mu)\Phi_o(t_{s,2\mu})$	$p_1(\frac{\lambda}{s+\lambda+\mu})(\frac{2\mu}{s+2\mu+\lambda})$
1,2	1	$t_{ba_1}+t_{ba_2}+t_{s,3\mu}$	$p_1\Psi_o(t_{ba_1}\|\mu)\Psi_o(t_{ba_2}\|2\mu)\Phi_o(t_{s,3\mu})$	$p_1(\frac{\lambda}{s+\lambda+\mu})(\frac{\lambda}{s+\lambda+2\mu})(\frac{3\mu}{s+3\mu+\lambda})$
.		.	.	.
.		.	.	.
.		.	.	.
1,(c-1)	1	$(\sum_{x=1}^{c-2} t_{ba_x})+t_{s,(c-1)\mu}$	$p_1[\prod_{x=2}^{c-2} \Psi_o(t_{ba_x}\|(x-1)\mu)]\Phi_o(t_{s,(c-1)\mu}$	$p_1(\frac{\lambda}{s+\lambda+\mu})\cdots(\frac{\lambda}{s+\lambda+(c-2)\mu})(\frac{(c-1)\mu}{s+(c-1)\mu+\lambda})$
1,c	1	$(\sum_{x=1}^{c-1} t_{ba_x})+t_{s,c\mu}$	$p_1[\prod_{x=2}^{c-1} \Psi_o(t_{ba_x}\|(x-1)\mu)]$	$p_1(\frac{\lambda}{s+\lambda+\mu})\cdots(\frac{\lambda}{s+\lambda+(c-1)\mu})(\frac{c\mu}{s+c\mu})$
2,0	2	$t_{s,2\mu}$	$p_2\Phi_o(t_{s,2\mu})$	$p_2(\frac{2\mu}{s+2\mu+\lambda})$
.		.	.	.
.		.	.	.

Table 2. Formulation of Output Distribution for M|M|c (Continued)

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
(c-2),0	c-2	$t_{s,(c-2)\mu}$	$p_{c-2}\Phi_o(t_{s,(c-2)\mu})$	$p_{c-2}(\frac{(c-2)\mu}{s+(c-2)\mu+\lambda})$
(c-2),1	c-2	$t_{ba}+t_{s,(c-1)\mu}$	$p_{c-2}\Psi_o(t_{ba}\|(c-2)\mu)\Phi_o(t_{s,(c-1)\mu})$	$p_{c-2}(\frac{\lambda}{s+\lambda+(c-2)\mu})(\frac{(c-1)\mu}{s+(c-1)\mu+\lambda})$
(c-2),2	c-2	$t_{ba_1}+t_{ba_2}+t_{s,c\mu}$	$p_{c-2}\Psi_o(t_{ba_1}\|(c-2)\mu)\Psi_o(t_{ba_2}\|(c-1)\mu)$	$p_{c-2}(\frac{\lambda}{s+\lambda+(c-2)\mu})(\frac{\lambda}{s+\lambda+(c-1)\mu})(\frac{c\mu}{s+c\mu})$
(c-1),0	c-1	$t_{s,(c-1)\mu}$	$p_{c-1}\Phi_o(t_{s,(c-1)\mu})$	$p_{c-1}(\frac{(c-1)\mu}{s+(c-1)\mu+\lambda})$

Table 2. Formulation of Output Distribution for M|M|c (Continued)

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
(c-1),1	c-1	$t_{ba} + t_{s,c\mu}$	$p_{c-1}\Psi_o(t_{ba} \mid (c-1)\mu)$	$p_{c-1}(\frac{\lambda}{s+\lambda+(c-1)\mu})(\frac{c\mu}{s+c\mu})$
c,0	$\geq c$	$t_{s,c\mu}$	$P_{\geq c}$	$P_{\geq c}(\frac{c\mu}{s+c\mu})$

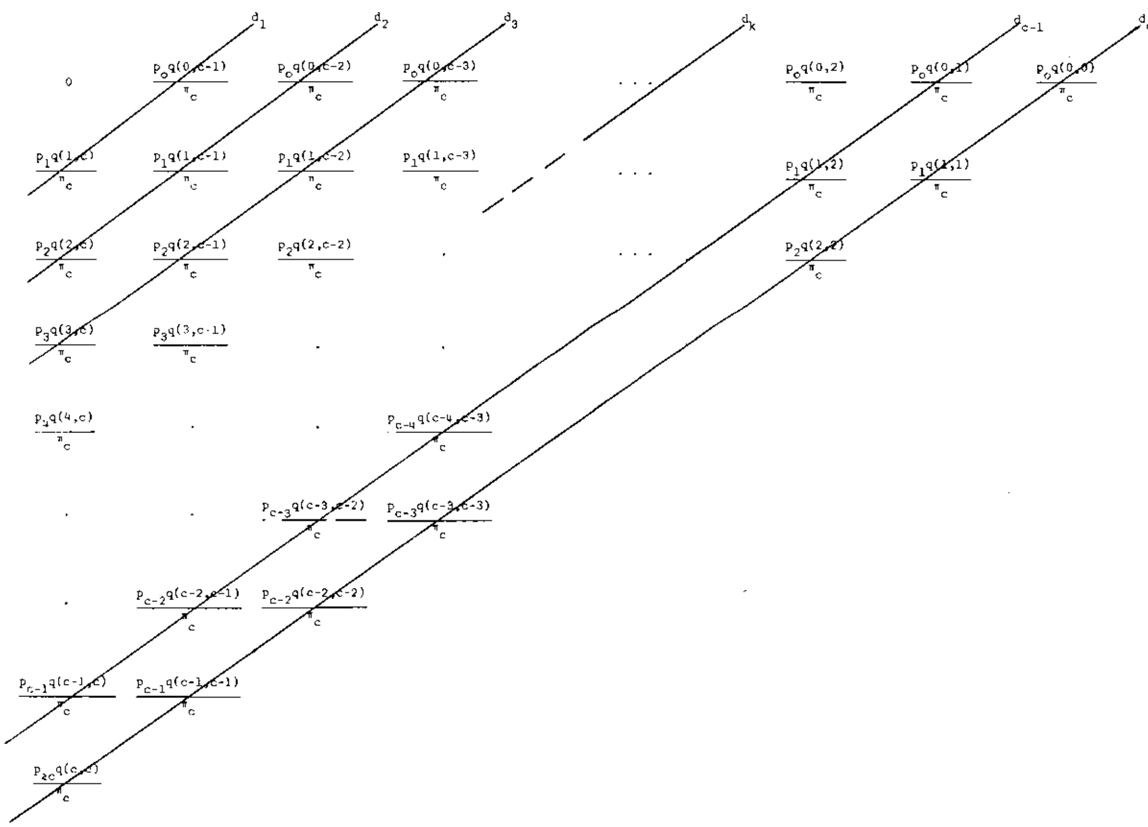


Figure 7. Terms in $h^e(s)$ for $M|M|c$.

Now sum the terms on the k^{th} diagonal $(d_k, k=1,\cdots,c-1)$:

$$d_k = \frac{1}{\pi_c}[p_k q(k,c) + p_{k-1}q(k-1,c-1) + \cdots + p_1 q(1,c-k+1) + p_o q(0,c-k)] \quad (58)$$

$$= \frac{1}{\pi_c}\{p_k[k\mu(x_o\cdots x_{k-1}x_{k+1}\cdots x_c)] + p_{k-1}[k\mu\lambda(x_o\cdots x_{k-2}x_{k+1}\cdots x_c]$$

$$+ \cdots + p_1[k\mu\lambda^{k-1}(x_o x_{k+1}\cdots x_c)]$$

$$+ p_o[k\mu\lambda^k(x_{k+1}\cdots x_c)] .$$

Rearranging terms,

$$d_k = \frac{k\mu(x_{k+1}\cdots x_c)}{\pi_c}[p_o\lambda^k + p_1 x_o\lambda^{k-1} + p_2 x_o x_1\lambda^{k-2} + \cdots \quad (59)$$

$$+ p_{k-2}(x_o\cdots x_{k-3})\lambda^2 + p_{k-1}(x_o\cdots x_{k-2})\lambda + p_k(x_o\cdots x_{k-1})] .$$

Considering the partial sums,

$$\xi_o = p_o\lambda^k , \quad (60)$$

$$\xi_1 = p_o\lambda^k + p_1 x_o\lambda^{k-1} = \frac{\lambda^k}{\mu} p_o x_1 ,$$

$$\xi_2 = p_o\lambda^k + p_1 x_o\lambda^{k-1} + p_2 x_o x_1\lambda^{k-2} = \frac{\lambda^{k-1}}{2\mu} p_1 x_1 x_2 ,$$

etc.,

we hypothesize that the partial sum,

$$\xi_m = p_o\lambda^k + p_1x_o\lambda^{k-1} + \cdots + p_m(x_o\cdots x_{m-1})\lambda^{k-m} , \tag{61}$$

is given by

$$\xi_m = \frac{\lambda^{k-m+1}}{m\mu}(x_1\cdots x_m)p_{m-1} , \quad m=1,2,\cdots,k . \tag{62}$$

The hypothesis has already been shown valid for m=1. Now consider

$$\xi_{m+1} = \xi_m + (x_o\cdots x_m)\lambda^{k-m-1}p_{m+1} \tag{63}$$

$$\overset{\text{set}}{=} \frac{\lambda^{k-m+1}}{m\mu}(x_1\cdots x_m)p_{m-1} + (x_o\cdots x_m)\lambda^{k-m-1}p_{m+1}$$

$$= (x_1\cdots x_m)\lambda^{k-m}\left[p_m + (x_{m+1}-(m+1)\mu)\frac{1}{\lambda}\frac{\lambda}{(m+1)\mu}p_m\right]$$

$$= \frac{\lambda^{k-m}}{(m+1)\mu}(x_1\cdots x_{m+1})p_m .$$

Thus, by induction, Equation (62) holds for $m=1,2,\cdots,k$.

In particular,

$$\xi_k = (x_1\cdots x_k)\frac{\lambda}{k\mu}p_{k-1} , \tag{64}$$

so that

$$d_k = \frac{k\mu(x_{k+1}\cdots x_c)}{\pi_c}(\xi_k) \tag{65}$$

$$= \frac{\pi_{c-1}}{\pi_c} \lambda p_{k-1}$$

$$= (\frac{\lambda}{s+\lambda}) p_{k-1} , \quad k=1,2,\cdots,c-1 .$$

For the remaining diagonal (d_c), we have

$$d_c = \frac{1}{\pi_c} [c\mu\lambda^c p_o + c\mu\lambda^{c-1} x_o p_1 + c\mu\lambda^{c-2} x_o x_1 p_2 + \cdots \tag{66}$$

$$+ c\mu(x_o \cdots x_{c-2}) p_{c-1} + c\mu(x_o \cdots x_{c-1}) P_{\geq c}]$$

$$= \frac{c\mu\lambda}{\pi_c} [\lambda^{c-1} p_o + \lambda^{c-2} x_o p_1 + \lambda^{c-3} x_o x_1 p_2 + \cdots$$

$$+ \lambda(x_o \cdots x_{c-3}) p_{c-2} + (x_o \cdots x_{c-2}) p_{c-1}]$$

$$+ \frac{c\mu}{\pi_c} (x_o \cdots x_{c-1}) P_{\geq c} .$$

The term inside the brackets in Equation (66) is our familiar partial sum, ξ_m, for m=c-1. Hence d_c reduces to

$$d_c = \frac{c\mu\lambda}{\pi_c} [(x_1 \cdots x_{c-1}) \frac{\lambda}{(c-1)\mu} p_{c-2}] + \frac{c\mu}{\pi_c} (x_o \cdots x_{c-1}) P_{\geq c} \tag{67}$$

$$= \frac{c\lambda^2}{(c-1)(x_o x_c)} p_{c-2} + \frac{c\mu}{x_c} P_{\geq c}$$

$$= \frac{c\lambda\mu}{(s+\lambda)(s+c\mu)} p_{c-1} + \frac{c\mu}{s+c\mu} P_{\geq c}$$

$$= \frac{\lambda}{s+\lambda} [p_{c-1} + P_{\geq c}] + \frac{s}{(s+\lambda)(s+c\mu)} [(c\mu-\lambda)P_{\geq c} - \lambda p_{c-1}] .$$

Combining the results of Equations (57), (65), and (67), we have

$$h^e(s) = \sum_{k=1}^{c-1} d_k + d_c \tag{68}$$

$$= (\frac{\lambda}{s+\lambda}) [\sum_{n=0}^{c-1} p_n + P_{\geq c}] + \frac{s}{(s+\lambda)(s+c\mu)} [(c\mu-\lambda)P_{\geq c} - \lambda p_{c-1}]$$

$$= \frac{\lambda}{s+\lambda} + \frac{s}{(s+\lambda)(s+c\mu)} [(c\mu-\lambda)P_{\geq c} - \lambda p_{c-1}] .$$

Equation (68) will prove useful in the derivation of an output distribution for the models which follow.

For the case at hand, we obtain upon substituting the values of p_{c-1} and $P_{\geq c}$ into Equation (68)

$$h^e(s) = \frac{\lambda}{s+\lambda} + \frac{s}{(s+\lambda)(s+c\mu)} [(c\mu-\lambda) \frac{(\lambda/\mu)^c p_o}{c!(1-\lambda/c\mu)} - \frac{\lambda(\lambda/\mu)^{c-1} p_o}{(c-1)!}] \tag{69}$$

$$= \frac{\lambda}{s+\lambda} + \frac{s\ p_o(\lambda/\mu)^{c-1}}{(c-1)!(s+\lambda)(s+c\mu)} [c\mu \frac{(\lambda/\mu)}{c} - \lambda]$$

$$= \frac{\lambda}{s+\lambda} .$$

Taking the inverse transform,

$$h(t_{bd}) = \lambda e^{-\lambda t_{bd}} , \quad t_{bd} \geq 0 , \tag{70}$$

completely independent of c and the servicing parameters.

Equation (70) is the desired expression for the output of the M|M|c model. The number of departures in time t is Poisson with the same parameter as the (Poisson) input distribution.

Output of the Truncated M|M|c Model

The truncated M|M|c model is a variation of the standard M|M|c model in which the queue is truncated at some point N-c; that is, values of n exceeding N are not allowed. The model is completely defined by specification of the λ_n and μ_n as

$$\left.\begin{array}{lll} n = 0 & \lambda_o = \lambda & \mu_o = 0 \\ n = 1,2,\cdots,c-1 & \lambda_n = \lambda & \mu_n = n\mu \\ n = c,c+1,\cdots,N-1 & \lambda_n = \lambda & \mu_n = c\mu \\ n = N & \lambda_N = 0 & \mu_N = c\mu \end{array}\right\} \tag{71}$$

Note that N has tacitly been assumed to be greater than or equal to c. If N were less than c, then c-N of the service stations would never be used. When these extraneous stations are eliminated from the model, it reduces to the case N=c.

The steady-state probabilities are obtained from Equation (32a);

$$p_n = \begin{cases} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n p_o & , \quad n=1,2,\cdots,c-1 ; \\ \\ \frac{1}{c!c^{n-c}}\left(\frac{\lambda}{\mu}\right)^n p_o & , \quad n=c,c+1,\cdots,N . \end{cases} \tag{72}$$

Again, for non-saturated queues, we require

$$\frac{\lambda}{c\mu} < 1 . \tag{73}$$

From Equation (35), we obtain

$$p_o = \left[\sum_{n=0}^{c-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^c \frac{1-\left(\frac{\lambda}{c\mu}\right)^{N-c+1}}{1-\left(\frac{\lambda}{c\mu}\right)}\right]^{-1} . \tag{74}$$

Since Equations (37) and (72) are identical, and since $\Phi_o(t)$, $\Psi_o(t|\mu_n)$, λ_n, and μ_n are the same in both the truncated and non-truncated cases for $n<c$, the previous derivation is applicable here if changes are made for p_o and $P_{\geq c}$. Note that a simple substitution of p_o and $P_{\geq c}$ for the truncated case will not suffice since a departee cannot leave N units in the system. (The amended Poisson assumptions do not allow an arrival and a departure to occur at the same epoch of time.) Hence we need the conditional probabilities

$$r_n \doteq \Pr(n|n\neq N) = \frac{p_n}{1-p_N}\ , \quad n=0,1,\cdots,N-1\ , \tag{75}$$

and the conditional cumulative probability,

$$R_{\geq c} \doteq \Pr(n\geq c|n\neq N) = \begin{cases} 0 & , N = c\ ; \\ \dfrac{P_{\geq c} - p_N}{1 - p_N} & , N > c\ , \end{cases} \tag{76}$$

where the p_n are given by Equations (72) and (74), to determine the output of the truncated M|M|c model.

Performing the indicated calculations, we can write

$$r_n = \begin{cases} \dfrac{1}{n!}\left(\dfrac{\lambda}{\mu}\right)^n r_o & , \quad n=1,2,\cdots,c-1\ ; \\ \dfrac{1}{c!c^{n-c}}\left(\dfrac{\lambda}{\mu}\right)^n r_o & , \quad n=c,c+1,\cdots,N-1 \text{ (if } N\neq c);\\ 0 & , \quad n=c \text{ (if } N=c)\ , \end{cases} \tag{77}$$

where,

$$r_o = \left[\sum_{n=0}^{c-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^c \left(\frac{1-\left(\frac{\lambda}{c\mu}\right)^{N-c}}{1-\left(\frac{\lambda}{c\mu}\right)}\right)\right]^{-1} . \tag{78}$$

Since,

$$P_{\geq c} = \frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^{c}\left(\frac{1-\left(\frac{\lambda}{c\mu}\right)^{N-c+1}}{1-\left(\frac{\lambda}{c\mu}\right)}\right) p_o \ , \tag{79}$$

then,

$$R_{\geq c} = \begin{cases} \frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^{c}\left(\frac{1-\left(\frac{\lambda}{c\mu}\right)^{N-c}}{1-\left(\frac{\lambda}{c\mu}\right)}\right) r_o \ , & N \neq c \ ; \\ 0 & , N=c \ . \end{cases} \tag{80}$$

Prior to Equation (69) in the derivation for the standard M|M|c model, no value of p_n was substituted. Further, the only relationship between the p_n which was used was

$$p_n = \left(\frac{\lambda_{n-1}}{\mu_n}\right)p_{n-1} = \left(\frac{\lambda}{\mu_n}\right)p_{n-1} \tag{81}$$

(which follows from Equation (32a)) for values of n less than c. Since this relationship holds also for the r_n (as does Equation 32), we have upon substituting $R_{\geq c}$ and r_{c-1} into Equation (68),

$$h_T^e(s) = \frac{\lambda}{s+\lambda} + \frac{s}{(s+\lambda)(s+c\mu)}\left[(c\mu-\lambda)R_{\geq c} - \lambda r_{c-1}\right] \ , \tag{82}$$

where the T on $h_T^e(s)$ indicates the truncated case.

Substituting from Equations (77) and (80), we obtain for $N \neq c$,

$$h_T^e(s) = \frac{\lambda}{s+\lambda} + \frac{s}{(s+\lambda)(s+c\mu)}\left[(c\mu-\lambda)\frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^c \left(\frac{1-\left(\frac{\lambda}{c\mu}\right)^{N-c}}{1-\left(\frac{\lambda}{c\mu}\right)}\right) r_o\right.$$

$$\left. -(\lambda)\frac{1}{(c-1)!}\left(\frac{\lambda}{\mu}\right)^{c-1} r_o\right]$$

$$= \frac{\lambda}{s+\lambda} - \left(\frac{s}{(s+\lambda)(s+c\mu)}\right)\left(\frac{\lambda}{(c-1)!}\left(\frac{\lambda}{\mu}\right)^{c-1}\left(\frac{\lambda}{c\mu}\right)^{N-c}\right) r_o \ .$$

For $N=c$, $R_{\geq c}=0$ and Equation (82) becomes

$$h_T^e(s) = \frac{\lambda}{s+\lambda} - \left(\frac{s}{(s+\lambda)(s+c\mu)}\right)\left(\frac{\lambda}{(c-1)!}\left(\frac{\lambda}{\mu}\right)^{c-1}\right) r_o \ .$$

Hence,

$$h_T^e(s) = \frac{\lambda}{s+\lambda} - \left(\frac{s}{(s+\lambda)(s+c\mu)}\right)\left(\frac{\lambda}{(c-1)!}\left(\frac{\lambda}{\mu}\right)^{c-1}\right)\left[\delta_{c,N} + (1-\delta_{c,N})\left(\frac{\lambda}{c\mu}\right)^{N-c}\right] r_o, \quad (83)$$

where r_o is defined in Equation (78) and $\delta_{c,N}$ is the Kronecker delta of order 2,

$$\delta_{i,j} = \begin{cases} 1 \ , & i=j \ ; \\ 0 \ , & i \neq j \ . \end{cases} \quad (84)$$

We can use a partial fraction to write $h_T^e(s)$ in the form,

$$h_T^e(s) = \frac{\lambda}{s+\lambda} - \{[(\frac{1}{c\mu-\lambda})(\frac{c\mu}{s+c\mu} - \frac{\lambda}{s+\lambda})][\frac{\lambda}{(c-1)!}(\frac{\lambda}{\mu})^{c-1}] \tag{85}$$

$$\cdot[\delta_{c,N} + (1-\delta_{c,N})(\frac{\lambda}{c\mu})^{N-c}]\, r_o\} \,.$$

Taking the inverse transform,

$$h_T(t_{bd}) = \lambda e^{-\lambda t_{bd}} - \{[(\frac{1}{c\mu-\lambda})(c\mu e^{-c\mu t_{bd}} - \lambda e^{-\lambda t_{bd}})] \tag{86}$$

$$\cdot[\frac{\lambda}{(c-1)!}(\frac{\lambda}{\mu})^{c-1}][\delta_{c,N} + (1-\delta_{c,N})(\frac{\lambda}{c\mu})^{N-c}]\, r_o\} \,,\ t_{bd} \geq 0 \,,$$

where r_o is defined in Equation (78).

Comparison of Equations (70) and (86) shows the effect of truncation on the time between departures. The mean time between departures for the truncated case is

$$E_T(t_{bd}) = \int_0^\infty t_{bd}\, h_T(t_{bd})dt_{bd} \tag{87}$$

$$= \frac{1}{\lambda}\{1 + [\frac{1}{c!}(\frac{\lambda}{\mu})^c][\delta_{c,N} + (1-\delta_{c,N})(\frac{\lambda}{c\mu})^{N-c}]\, r_o\}$$

and the mean number of arrivals lost per unit time is just the difference between the mean arrival and departure rates,

$$[\lambda - E_T(t_{bd})^{-1}] = \lambda - \lambda\{1 + [\frac{1}{c!}(\frac{\lambda}{\mu})^c][\delta_{c,N} + (1-\delta_{c,N})(\frac{\lambda}{c\mu})^{N-c}]\, r_o\}^{-1} \tag{88}$$

$$= p_N\lambda \,.$$

The result of Equation (88) should have been anticipated from intuitive arguments.

Output of the Queues-with-Discouragement Model

The queues-with-discouragement model is an $M_n|M|c$ model in which new arrivals are discouraged in direct proportion to their expected wait for service. The model is defined by

$$\left.\begin{array}{lll} n = 0 & \lambda_n = \lambda & \mu_o = 0 \\ n = 1,2,\cdots,c-1 & \lambda_n = \lambda & \mu_n = n\mu \\ n = xc, xc+1,\cdots,(x+1)c-1 & \lambda_n = \lambda/x & \mu_n = c\mu \\ \qquad (x=1,2,\cdots) & & \end{array}\right\} \tag{89}$$

where the proper value of x to be substituted into the expression for λ_n is obtained as the largest integer in (n/c). From Equation (32a), we obtain the steady-state probabilities,

$$p_n = \begin{cases} \dfrac{1}{n!}\left(\dfrac{\lambda}{\mu}\right)^n p_o \ , & n=1,2,\cdots,c-1 \ ; \\ \dfrac{1}{(x!)^c (x+1)^{n-xc} c! c^{n-c}}\left(\dfrac{\lambda}{\mu}\right)^n p_o \ , & n=xc,xc+1,\cdots,(x+1)c-1 \ ; \\ & x=1,2,\cdots \ , \end{cases} \tag{90}$$

where p_o is defined by Equation (35) as

$$p_o = [\sum_{n=0}^{c-1} \frac{1}{n!} (\frac{\lambda}{\mu})^n + \sum_{x=1}^{\infty} \sum_{n=xc}^{(x+1)c-1} \frac{1}{(x!)^c (x+1)^{n-xc} c! c^{n-c}} (\frac{\lambda}{\mu})^n]^{-1} . \quad (91)$$

$P_{\geq c}$ is determined by summing the p_n from c to ∞ , that is,

$$P_{\geq c} = \sum_{n=c}^{\infty} p_n = \sum_{x=1}^{\infty} \sum_{n=xc}^{(x+1)c-1} \frac{1}{(x!)^c (x+1)^{n-xc} c! c^{n-c}} (\frac{\lambda}{\mu})^n p_o . \quad (92)$$

Since the p_n were determined from Equation (32a) and since $\lambda_n = \lambda$ for all n, Equation (81) can be seen to hold for this model. Thus, Equation (68) holds and

$$h_D^e(s) = \frac{\lambda}{s+\lambda} + \frac{s}{(s+\lambda)(s+c\mu)} [(c\mu-\lambda)P_{\geq c} - \lambda p_{c-1}] \quad (93)$$

$$= \frac{\lambda}{s+\lambda} + [(\frac{1}{c\mu-\lambda})(\frac{c\mu}{s+c\mu} - \frac{\lambda}{s+\lambda})][(c\mu-\lambda)P_{\geq c} - \lambda p_{c-1}]$$

$$= \frac{\lambda}{s+\lambda} + [(\frac{1}{c\mu-\lambda})(\frac{c\mu}{s+c\mu} - \frac{\lambda}{s+\lambda})]$$

$$\cdot [(c\mu-\lambda) \sum_{x=1}^{\infty} \sum_{n=xc}^{(x+1)c-1} \frac{1}{(x!)^c (x+1)^{n-xc} c! c^{n-c}} (\frac{\lambda}{\mu})^n p_o$$

$$- \lambda \frac{1}{(c-1)!} (\frac{\lambda}{\mu})^{c-1} p_o]$$

$$= \frac{\lambda}{s+\lambda} - [\frac{c\mu}{s+c\mu} - \frac{\lambda}{s+\lambda}][\frac{1}{c!} (\frac{\lambda}{\mu})^c p_o][\frac{1}{1 - \lambda/(c\mu)}$$

$$- \sum_{x=1}^{\infty} \sum_{n=xc}^{(x+1)c-1} (\frac{1}{(x!)^c (x+1)^{n-xc}} (\frac{\lambda}{\mu})^{n-c})] ,$$

where the D on $h_D^e(s)$ indicates the queues-with-discouragement model.

Taking the inverse transform of Equation (93), we obtain the desired result,

$$h_D(t_{bd}) = \lambda e^{-\lambda t_{bd}} - [c\mu e^{-c\mu t_{bd}} - \lambda e^{-\mu t_{bd}}] \tag{94}$$

$$\cdot[\frac{1}{c!}(\frac{\lambda}{\mu})^c p_o][\frac{1}{1-\lambda/(c\mu)} - \sum_{x=1}^{\infty}\sum_{n=xc}^{(x+1)c-1}(\frac{1}{(x!)^c(x+1)^{n-xc}c^{n-c}}(\frac{\lambda}{\mu})^{n-c})] ,$$

where p_o is defined by Equation (91).

Comparison of Equations (70) and (94) shows the effect of discouragement on the time between departures. The mean time between departures for the discouraged case is

$$E_D(t_{bd}) \doteq \int_o^{\infty} t_{bd}\, h_D(t_{bd})dt_{bd} \tag{95}$$

$$= \frac{1}{\lambda}\{1 + [\frac{1}{c!}(\frac{\lambda}{\mu})^c p_o][1 - \sum_{x=1}^{\infty}\sum_{n=xc}^{(x+1)c-1}(\frac{1-\lambda/(c\mu)}{(x!)^c(x+1)^{n-xc}c^{n-c}}(\frac{\lambda}{\mu})^{n-c})]\}$$

and the mean number of arrivals lost per unit time is

$$[\lambda - E_D(t_{bd})^{-1}] = \{\frac{1}{c!}(\frac{\lambda}{\mu})^c p_o \tag{96}$$

$$\cdot[\frac{1}{\lambda} - \sum_{x=1}^{\infty}\sum_{n=xc}^{(x+1)c-1}(\frac{1-\lambda/(c\mu)}{(x!)^c(x+1)^{n-xc}c^{n-c}}(\frac{\lambda}{\mu})^{n-c})]\} .$$

Output of the General $M_n|M_n|1$ Model

The three $M_n|M|c$ models treated above comprise the major portion of $M_n|M_n|1$ models in current use.* However, since there is an infinite number of other possible models, it might prove useful to have an expression for the output of the general $M_n|M_n|1$. This expression, Equation (99), is offered more for the sake of completeness than for any intended application. It is suggested that, in all but the simplest cases, an output distribution can be obtained more quickly and in a more concise form by setting up an interdeparture event table similar to those presented in this chapter and studying the combinatorial relationships of a few terms at a time.

The steady-state probabilities of the general $M_n|M_n|1$ model are given by Equations (32a) and (35). With the exception of the physical constraint,

$$\mu_o = 0 \ , \tag{97}$$

values of λ_n and μ_n may be independently specified, or may be related by any function of n whatsoever.

The output distribution is formulated in exactly the same way as for the previously discussed models. The results of this formulation

---

*Remember that the $M_n|M_n|1$ model is only tacitly assumed to possess but a single input source and a single channel. Multi-input and multichannel models are obtainable through suitable specification and interpretation of relationships among the λ_n and μ_n. The $M_n|M_n|1$ is the most general of those with Poisson input and negative exponential service time distributions.

are shown in Table 3. The transformed density function for the time between departures, $h^e(s)$, is the sum of the terms in the last column. Thus,

$$h^e(s) = \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \{p_n[\delta_{o,k} + (1-\delta_{o,k}) \prod_{x=1}^{k} (\frac{\lambda_{n+x-1}}{s+\lambda_{n+x-1}+\mu_{n+x-1}})] \quad (98)$$

$$\cdot[\frac{\mu_{n+k}}{s+\mu_{n+k}+\lambda_{n+k}}]\} \ ,$$

where $\delta_{o,k}$ is the Kronecker delta defined in Equation (84). Substituting from Equation (32a), we obtain the final form,

$$h^e(s) = p_o \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \{[\delta_{o,n} + (1-\delta_{o,n}) \prod_{y=1}^{n} (\frac{\lambda_{y-1}}{\mu_y})] \quad (99)$$

$$\cdot[\delta_{o,k} + (1-\delta_{o,k}) \prod_{x=1}^{k} (\frac{\lambda_{n+x-1}}{s+\lambda_{n+x-1}+\mu_{n+x-1}})][\frac{\mu_{n+k}}{s+\mu_{n+k}+\lambda_{n+k}}]\} \ ,$$

where p_o is defined by Equation (35).

Two observations should be made about Table 3 and Equation (99). First it can be seen that the finding of an inverse transform for Equation (99) is a tedious process. In some cases a solution might be obtained from the inverse Laplace transform equation,

$$h(t_{bd}) = \mathcal{L}^{-1}[h^e(s)] = \frac{1}{2\pi i} \int_{z-i\infty}^{z+i\infty} e^{st} h^e(s) ds \ , \quad (100)$$

where z is chosen to the right of any singularity of $h^e(s)$, or from partial fraction expansion of the second product expression and subsequent term by term inversion using standard tables. In other cases, the complexity of $h^e(s)$ might be reduced by a judicious combination of terms as was done in the reduction centered around Figure 7.*

Second, care must be exercised in the writing of Table 3 for the allowed values of n. All terms involving p_i where p_i is always zero (via Equation (32a)) may be summarily dropped. However, if some p_i is zero only at the departure epoch (as was the case with p_N in the truncated M|M|c model), then the terms involving p_i are dropped <u>and</u> the other p_n are increased accordingly ($p_n \rightarrow p_n/(1-p_i)$). Equation (98) holds for either Equation (32a) or Equation (32b) and thus, if some μ_n for $n \geq 1$ is zero, an alternative to Equation (99) can be obtained by substituting Equation (32b) into Equation (98).

Third, care must also be exercised in the writing of Table 3 for the allowed values of k. Most important, of course, is to be sure to include all the possible events which lead to the next departure epoch and consequently define t_{bd}. It is suggested that sets of events be chosen such that their intersection is empty. This eliminates the problem of subtracting the joint probabilities of the subevents when summing to obtain $h(t_{bd})$ or $h^e(s)$ for the union of all the sets of events. A set of drawings similar to Figure 6 will probably be helpful in this regard.

---

*Note that an inverse transformation may always be accomplished if one does not insist on an explicit solution for $h(t_{bd})$. Consider multiplying both sides of Equation (99) by the lowest common denominator (l.c.d.) to obtain (l.c.d.)$h^e(s)$ equal to a polynomial in s. The inverse transform of this equation is well-known (standard tables) to be a time domain differential equation in $h(t_{bd})$.

Table 3. Formulation of Output Distribution for $M_n \mid M_n \mid 1$

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
0,1	0	$t_{ba,\lambda_o}+t_{s,\mu_1}$	$p_o\Phi_o(t_s \mid \lambda_1,\mu_1)$	$p_o(\frac{\lambda_o}{s+\lambda_o})(\frac{\mu_1}{s+\mu_1+\lambda_1})$
0,2	0	$t_{ba_1,\lambda_o}+t_{ba_2,\lambda_1}+t_{s,\mu_2}$	$p_o\Psi_o(t_{ba_2} \mid \lambda_1,\mu_1)\Phi_o(t_s \mid \lambda_2,\mu_2)$	$p_o(\frac{\lambda_o}{s+\lambda_o})(\frac{\lambda_1}{s+\lambda_1+\mu_1})(\frac{\mu_2}{s+\mu_2+\lambda_2})$
0,3	0	$t_{ba_1,\lambda_o}+t_{ba_2,\lambda_1}+t_{ba_3,\lambda_2}+t_{s,\mu_3}$	$p_o\Psi_o(t_{ba_2} \mid \lambda_1,\mu_1)\Psi_o(t_{ba_3} \mid \lambda_2,\mu_2)$ $\cdot\,\Phi_o(t_s \mid \lambda_3,\mu_3)$	$p_o(\frac{\lambda_o}{s+\lambda_o})(\frac{\lambda_1}{s+\lambda_1+\mu_1})(\frac{\lambda_2}{s+\lambda_2+\mu_2})(\frac{\mu_3}{s+\mu_3+\lambda_3})$
⋮	⋮	⋮	⋮	⋮
0,k	0	$(\sum_{x=1}^{k} t_{ba_x,\lambda_{x-1}})+t_{s,\mu_k}$	$p_o[\prod_{x=1}^{k}\Psi_o(t_{ba_x} \mid \lambda_{x-1},\mu_{x-1})]$ $\cdot\,[\Phi_o(t_s \mid \lambda_k,\mu_k)]$	$p_o[\prod_{x=1}^{k}\frac{\lambda_{x-1}}{s+\lambda_{x-1}+\mu_{x+1}}](\frac{\mu_k}{s+\mu_k+\lambda_k})$

Table 3. Formulation of Output Distribution for $M_n|M_n|1$ (Continued)

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
⋮	⋮	⋮	⋮	⋮
0,∞				
1,0	1	t_{s,μ_1}	$p_1\Phi_o(t_s\|\lambda_1,\mu_1)$	$p_1(\frac{\mu_1}{s+\mu_1+\lambda_1})$
1,1	1	$t_{ba,\lambda_1}+t_{s,\mu_2}$	$p_1\Psi_o(t_{ba}\|\lambda_1,\mu_1)\Phi_o(t_s\|\lambda_2,\mu_2)$	$p_1(\frac{\lambda_1}{s+\lambda_1+\mu_1})(\frac{\mu_2}{s+\lambda_2+\mu_2})$
1,2	1	$t_{ba_1,\lambda_1}+t_{ba_2,\lambda_2}+t_{s,\mu_3}$	$p_1\Psi_o(t_{ba_1}\|\lambda_1,\mu_1)\Psi_o(t_{ba_2}\|\lambda_2,\mu_2)\Phi_o(t_s\|\lambda_3,\mu_3)$	$p_1(\frac{\lambda_1}{s+\lambda_1+\mu_1})(\frac{\lambda_2}{s+\lambda_2+\mu_2})(\frac{\mu_3}{s+\mu_3+\lambda_3})$
⋮	⋮	⋮	⋮	⋮

Table 3. Formulation of Output Distribution for $M_n|M_n|1$ (Continued)

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
.	.	.	.	.
1,k	1	$(\sum_{x=1}^{k} t_{ba_x,\lambda_x})+t_{s,\mu_k}$	$p_1[\prod_{x=1}^{k}\Psi_o(t_{ba_x}\|\lambda_x,\mu_x)]\Phi_o(t_s\|\lambda_{k+1},\mu_{k+1})$	$p_1[\prod_{x=1}^{k}\frac{\lambda x}{s+\lambda_x+\mu_x}](\frac{\mu_{k+1}}{s+\mu_{k+1}+\lambda_{k+1}})$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
1,∞				
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.

Table 3. Formulation of Output Distribution for $M_n|M_n|1$ (Continued)

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
j,0	j	t_{s,μ_j}	$p_j \Phi_o(t_s \mid \lambda_j,\mu_j)$	$p_j\left(\frac{\mu_j}{s+\mu_j+\lambda_j}\right)$
j,1	j	$t_{ba,\lambda_j}+t_{s,\mu_{j+1}}$	$p_j \Psi_o(t_{ba} \mid \lambda_j,\mu_j)\Phi_o(t_s \mid \lambda_{j+1},\mu_{j+1})$	$p_j\left(\frac{\lambda_j}{s+\lambda_j+\mu_j}\right)\left(\frac{\mu_{j+1}}{s+\mu_{j+1}+\lambda_{j+1}}\right)$
j,2	j	$t_{ba_1,\lambda_j}+t_{ba_2,\lambda_{j+1}}+t_{s,\mu_{j+2}}$	$p_j \Psi_o(t_{ba_1} \mid \lambda_j,\mu_j)\Psi_o(t_{ba_2} \mid \lambda_{j+1},\mu_{j+1})$	$p_j\left(\frac{\lambda_j}{s+\lambda_j+\mu_j}\right)\left(\frac{\lambda_{j+1}}{s+\lambda_{j+1}+\mu_{j+1}}\right)$
			$\cdot\ \Phi_o(t_s \mid \lambda_{j+2},\mu_{j+2})$	$\cdot\ \left(\frac{\mu_{j+2}}{s+\mu_{j+2}+\lambda_{j+2}}\right)$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
j,k	j	$\left(\sum_{x=1}^{k} t_{ba_x,\lambda_{j+x-1}}\right)+t_{s,\mu_{j+k}}$	$p_j\left[\prod_{x=1}^{k} \Psi_o(t_{ba_x} \mid \lambda_{j+x-1},\mu_{j+x-1})\right]$	$p_j\left[\prod_{x=1}^{k}\left(\frac{\lambda_{j+x-1}}{s+\lambda_{j+x-1}+\mu_{j+x-1}}\right)\right]$
			$\cdot\ [\Phi_o(t_s \mid \lambda_{j+k},\mu_{j+k})]$	$\cdot\ \left(\frac{\mu_{j+k}}{s+\mu_{j+k}+\lambda_{j+k}}\right)$

Table 3. Formulation of Output Distribution for $M_n|M_n|1$ (Continued)

Event	n	t_{bd}	With Probability	$p_n h^e(s)$
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
j,∞				
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
∞,∞				

CHAPTER IV

OUTPUT OF MORE GENERAL MODELS

G|G|c denotes those multichannel models which have a general class of input and service time distributions. If independence assumptions (i.e., inter-event intervals are identically and independently distributed by a general distribution) are made, the GI|GI|c models result. It is members of this latter group which will be studied in this chapter.

There are several reasons for restricting analysis of the general models to those with independence assumptions. One is the difficulty of formulating an approach with the limited information available on G|G|c models. Most of the literature on G|G|c queues deals with the single channel case, often giving only expected values of parameters rather than their frequency distributions. Multichannel results are generally restricted to studies of waiting times and the busy period.

Indeed, even the GI|GI|c literature is somewhat limited. For example, no general expression is known for the equilibrium state probabilities of the M|Gl|c model studied below.

Output of M|GI|c

Here we study the multichannel model with Poisson input and channel service times which are identically, independently, and arbitrarily distributed. The method of attack is a variation on the enumerative approach of Chapter III. The result obtained in the multichannel case will not be immediately useful since the steady-state probabilities, p_n, are not known. However, in the vein of Descamp [19] and Saaty [69] we will not let this be a deterent since, for application, the p_n might be determined from experiment or other approaches (e.g., [21]).

We will retain the previous notation and introduce new symbols as they are needed. In particular, recall that $h(t_{bd})$ denotes the interdeparture interval density function; $g(t_{ba})$, the interarrival interval density function; and $\Phi_o(t)$, the probability that no arrivals occur in time t. The symbol $b(t_s)$ will be used as the general independent service time density function; λ, as the mean (Poisson) arrival rate; and μ, as the mean service rate.

M|GI|1

The generating function of the M|GI|1 state probabilities is well known (e.g. [69], p. 194):

$$p_z^g = \frac{(1 - \frac{\lambda}{\mu})b^e[\lambda(1-z)]}{1 - \{1-b^e[\lambda(1-z)]\}/(1-z)} \quad , \quad \frac{\lambda}{\mu} < 1 \ . \tag{101}$$

To find the output distribution, suppose a departure has occurred at t = 0. Then, with probability $(1-p_o)$,

$$t_{bd} = t_s \quad \text{and} \quad h(t_{bd}) = b(t_{bd}) \ , \tag{102}$$

and, with probability p_o,

$$t_{bd} = t_{ba} + t_s \quad \text{and} \quad h(t_{bd}) = g(t_{bd}) * b(t_{bd}) \ , \tag{103}$$

where the Markovian property of the negative exponential distribution, $g(t_{ba})$, has been used to establish t_{ba} as the waiting time to the next arrival. Thus,

$$h^e(s) = [(1-p_o) + p_o(\frac{\lambda}{s+\lambda})]b^e(s) \ , \tag{104}$$

and $h(t_{bd})$ may be obtained by taking the inverse transform of Equation (104).

Equation (104) is a well-known result. The procedure may be used as an alternative to Chang's method [10], which was reported in Chapter II.

M|GI|c

Takács [75] and Descamp [19] have independently obtained the waiting time distribution for M|GI|1 and Descamp has extended the result to c channels. Their formulations required an expression for the remaining service time on a unit when observations are started at some arbitrary time t. This is just what we will need to write the probabilities of a set of mutually independent and collectively exhaustive events as we did in Chapter III.

Let $\gamma_o(t)$ be the probability that a channel which was occupied at the time of a departure (t=0), continues to be occupied after time t. Then,

$$\gamma_o(t) = \frac{1}{1/\mu} \int_t^{\infty} [1 - \int_o^{\tau} b(x)dx]d\tau \ , \tag{105}$$

where $1/\mu$ is the average service time of a unit. The expression in the numerator is just the expected time of occupancy of the channel after time t. The ratio gives the desired probability.

As a further elaboration, Saaty ([69] p. 204) notes that, necessarily,

$$\begin{aligned}
\gamma_o(0) &= \mu \int_o^\infty [1 - \int_o^\tau b(x)dx]d\tau \\
&= \mu \lim_{s\to o} \int_o^\infty e^{-s\tau} [1 - \int_o^\tau b(x)dx]d\tau \\
&= \mu \lim_{s\to o} [\frac{1}{s} - \int_o^\infty b(x)dx \int_x^\infty e^{-s\tau}d\tau] \\
&= \mu \lim_{s\to o} [\frac{1 - b^e(s)}{s}] \\
&= 1 ,
\end{aligned}$$

after applying L'Hospital's rule.

If we let $\Theta_o(t)$ denote the complementary cumulative distribution function of b(t), i.e.

$$\Theta_o(t) = 1 - \int_o^t b(x)dx , \tag{106}$$

we have

$$\gamma_o(t) = \frac{1}{1/\mu} \int_t^\infty \Theta_o(\tau)d\tau . \tag{107}$$

Combinations of $\gamma_o(t)$ and $\Theta_o(t)$ will take on the role played by $\Psi_o(t)$

in our previous formulation. In particular, note that the probability of zero service completions in time t after the last previous departure, if he left c or more units in the system, is

$$[\gamma_o(t)]^{c-1}\ \theta_o(t) \tag{108}$$

under the independence assumptions.

Also needed will be an expression for the probability that a service completion occurs at time t when observations are started at some arbitrary time t = 0. Denoting this probability by $\beta(t)$, we have

$$\beta(t) = \frac{d}{dt}\,[1 - \gamma_o(t)] = -\frac{d\gamma_o(t)}{dt}\ . \tag{109}$$

Now all that remains is to select a set of independent inter-departure events and formulate an expression for $h(t_{bd})$. We will use a set similar to those chosen in Chapter III. It will be convenient to classify the events as belonging to one of two groups. The first group includes those in which a new arrival (excluding the unit from the queue which replaces the departee) is the first to complete service. Define the events,

$$E(n,k,m) \doteq \left\{ \begin{array}{l} \text{n units in the system at } t=0^+\text{; k new arrivals} \\ \text{before a service completion; the } m^{th} \text{ new arrival} \\ \text{completes his service before any other unit} \end{array} \right\} ,$$

where, $n = 0,1,\cdots,c-1$;

$k = 1,2,\cdots,(c-n)^+$;

$$m = 1,2,\cdots, \min[k,c-n] \ ;$$

and $(c-n)^+$ indicates "c-n, or more."

In the second group are those events in which one of the units present at the time of the departure is the next to complete service. Define the events,

$$E(n,k,\cdot) \doteq \left\{ \begin{array}{l} \text{n units in the system at } t = 0^+\text{; k new arrivals} \\ \text{before a service completion; a unit in service at} \\ t = 0^+ \text{ is the first to complete service} \end{array} \right\} ,$$

where,
$$n = 0,1,\cdots,c+ \ ;$$
$$k = 0,1,\cdots,(c-n)^+ \ .$$

An algebraic expression can now be written for the time between departures for each event and the corresponding probabilities, $p_n h(t_{bd})$, can be determined as shown in Table 4. The desired density function, $h(t_{bd})$, for the model is obtained by summing the weighted event density functions as was done for the other tables. This yields

$$h(t_{bd}) = \sum_{n=0}^{c-1} \{p_n[\gamma_o(t_{bd})]^n \int_o^{t_{bd}} g(t_{bd}-x)b(x)\Phi_o(x)dx\} \qquad (110)$$

$$+ \sum_{n=0}^{c-2} \sum_{k=2}^{c-n} \sum_{m=1}^{k} \{ \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int_o^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Z_1)dxdt_{ba_k} \cdots dt_{ba_2}\}$$

$$+ \sum_{n=1}^{c-1} \{p_n \ \beta(t_{bd})[\gamma_o(t_{bd})]^{n-1} \ \Phi_o(t_{bd})\}$$

$$+\sum_{n=1}^{c-1}\{p_n\ \beta(t_{bd})[\gamma_o(t_{bd})]^{n-1}\int_o^{t_{bd}} g(t_{bd}-x)\Theta_o(x)\Phi_o(x)dx\}$$

$$+\sum_{n=1}^{c-2}\sum_{k=2}^{c-n}\{\int_o^{t_{bd}}\int_o^{t_{bd}-t_{ba_2}}\cdots\int_o^{t_{bd}-\sum_{j=2}^{k}t_{ba_j}}(Z_2)dxdt_{ba_k}\cdots dt_{ba_2}\}$$

$$+P_{\geq c}[\gamma_o(t_{bd})]^{c-1}\ b(t_{bd}) + P_{\geq c}[\gamma_o(t_{bd})]^{c-2}\ \Theta_o(t_{bd})\beta(t_{bd})\ ,$$

where

$$Z_1 = p_n[\gamma_o(t_{bd})]^n[g(t_{bd}-\sum_{j=2}^{k}t_{ba_j}-x)][\prod_{j=2}^{k}g(t_{ba_j})] \tag{111}$$

$$\cdot[\delta_{k,m}\ b(x) + (1-\delta_{k,m})\ b(\sum_{j=m+1}^{k}t_{ba_j}+x)]$$

$$\cdot[\prod_{\substack{i=2\\ i\neq m+1}}^{k}\Theta_o(\sum_{j=1}^{k}t_{ba_j}+x)][\delta_{k,m}+(1-\delta_{k,m})\Theta_o(x)]\Phi_o(x)\ ,$$

$$Z_2 = p_n[\gamma_o(t_{bd})]^{n-1}\beta(t_{bd})[g(t_{bd}-\sum_{j=2}^{k}t_{ba_j}-x)][\prod_{j=2}^{k}g(t_{ba_j})] \tag{112}$$

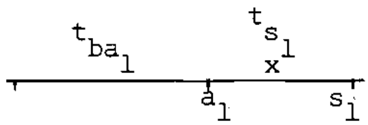
$$\cdot[\prod_{i=2}^{k}\Theta_o(\sum_{j=i}^{k}t_{ba_j}+x)]\Theta_o(x)\Phi_o(x)\ ,$$

and $\delta_{k,m}$ is the Kronecker delta of Equation (84).

Equation (110) is quite unwieldy and it is suggested that solutions be obtained in the form of Table 4 rather than by substituting directly in the equation. The remarks following Equation (99) also apply to the construction of Table 4 and to the possible simplification

Table 4. Formulation of Output Distribution for M|GI|c

Event (0,1,1)

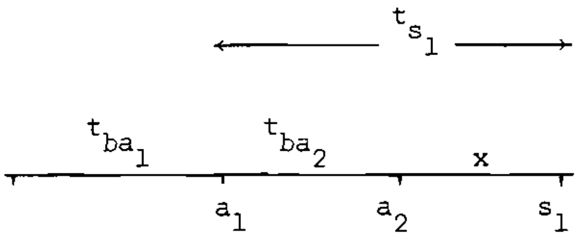


$$t_{bd} = t_{ba_1} + t_{s_1} = t_{ba_1} + x$$

with probability $p_o \Phi_o(x)$

$$p_o h(t_{bd}) = p_o \int_o^{t_{bd}} g(t_{bd}-x) b(x) \Phi_o(x) dx$$

Event (0,2,1)



$$t_{bd} = t_{ba_1} + t_{s_1} = t_{ba_1} + t_{ba_2} + x$$

with probability $p_o \Phi_o(x) \Theta_o(x)$

$$p_o h(t_{bd}) = p_o \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} g(t_{bd}-t_{ba_2}-x) g(t_{ba_2}) b(t_{ba_2}+x) \cdot \Theta_o(x) \Phi_o(x) dx dt_{ba_2}$$

Event (0,2,2)

$$t_{bd} = t_{ba_1} + t_{ba_2} + t_{s_2} = t_{ba_1} + t_{ba_2} + x$$

with probability $p_o \Theta_o(t_{ba_2} + x) \Phi_o(x)$

Table 4. Formulation of Output Distributions for M|GI|c (Continued)

$$p_o h(t_{bd}) = p_o \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} g(t_{bd}-t_{ba_2}-x)g(t_{ba_2})b(x)$$

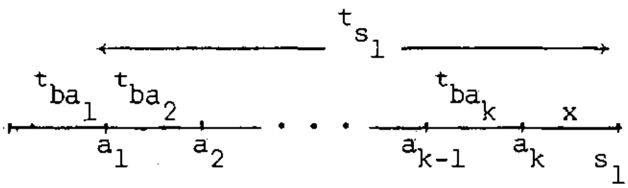
$$\cdot\ \Theta_o(t_{ba_2}+x)\Phi_o(x)dxdt_{ba_2}$$

.

.

.

Event (0,k,1)



$$t_{bd} = t_{ba_1} + t_{s_1} = \sum_{j=1}^{k} t_{ba_j} + x$$

$$\text{with probability } p_o\Big[\prod_{i=3}^{k} \Theta_o\Big(\sum_{j=i}^{k} t_{ba_j} + x\Big)\Big]\Theta_o(x)\Phi_o(x)$$

$$p_o h(t_{bd}) = \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int_o^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Z)dxdt_{ba_k}\cdots dt_{ba_2}\ ,$$

where

$$Z = p_o\Big[g\Big(t_{bd}-\sum_{j=2}^{k} t_{ba_j}-x\Big)\Big]\Big[\prod_{j=2}^{k} g(t_{ba_j})\Big]\Big[b\Big(\sum_{j=2}^{k} t_{ba_j}+x\Big)\Big]$$

$$\cdot\ \Big[\prod_{i=3}^{k} \Theta_o\Big(\sum_{j=i}^{k} t_{ba_j}+x\Big)\Big]\Theta_o(x)\Phi_o(x)$$

Table 4. Formulation of Output Distributions for M|GI|c (Continued)

$$\xleftarrow{\qquad\qquad} t_{s_2} \xrightarrow{\qquad\qquad}$$

<u>Event (0,k,2)</u>

$$\underset{\qquad a_1}{t_{ba_1}} \quad \underset{a_2}{t_{ba_2}} \quad \underset{a_3}{t_{ba_3}} \quad \cdots \quad \underset{a_{k-1}}{} \quad \underset{a_k}{t_{ba_k}} \quad \underset{s_2}{x}$$

$$t_{bd} = t_{ba_1} + t_{ba_2} + t_{s_2} = \sum_{j=1}^{k} t_{ba_j} + x$$

$$\text{with probability } p_o[\prod_{\substack{i=2 \\ i \neq 3}}^{k} \Theta_o(\sum_{j=1}^{k} t_{ba_j} + x)]\Theta_o(x)\Phi_o(x)$$

$$p_o h(t_{bd}) = \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int_o^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Z)dxdt_{ba_k} \cdots dt_{ba_2} \text{ ,}$$

where

$$Z = p_o[g(t_{bd} - \sum_{j=2}^{k} t_{ba_j} - x)] [\prod_{j=2}^{k} g(t_{ba_j})] [b(\sum_{j=3}^{k} t_{ba_j} + x)]$$

$$\cdot [\prod_{\substack{i=2 \\ i \neq 3}}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j} + x)] \Theta_o(x)\Phi_o(x)$$

.

.

.

Table 4. Formulation of Output Distribution for M|GI|c (Continued)

Event (0,k,m)

$$t_{bd} = \sum_{j=1}^{m} t_{ba_j} + t_{s_m} = \sum_{j=1}^{k} t_{ba_j} + x$$

$$\text{with probability } p_o[\prod_{\substack{i=2 \\ i \neq m+1}}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j} + x)]\Theta_o(x)\Phi_o(x)$$

$$p_o h(t_{bd}) = \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int_o^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Z)dxdt_{ba_k}\cdots dt_{ba_2} \ ,$$

where

$$Z = p_o[g(t_{bd} - \sum_{j=2}^{k} t_{ba_j} - x)][\prod_{j=2}^{k} g(t_{ba_j})][b(\sum_{j=m+1}^{k} t_{ba_j} + x)]$$

$$\cdot[\sum_{\substack{i=2 \\ i \neq m+1}}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j} + x)]\Theta_o(x)\Phi_o(x)$$

.

.

.

Event (0,k,k)

$$t_{bd} = \sum_{j=1}^{k} t_{ba_j} + t_{s_k} = \sum_{j=1}^{k} t_{ba_j} + x$$

Table 4. Formulation of Output Distribution for M|GI|c (Continued)

$$\text{with probability } p_o[\prod_{i=2}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j})]\Phi_o(x)$$

$$p_o h(t_{bd}) = \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int_o^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Z)dxdt_{ba_k} \cdots dt_{ba_2} \; ,$$

where $$Z = p_o[g(t_{bd}-\sum_{j=2}^{k} t_{ba_j} - x)][\prod_{j=2}^{k} g(t_{ba_j})][b(x)]$$

$$\cdot[\prod_{i=2}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j} + x)]\Phi_o(x)$$

.

.

.

Event $(0,c^+,1)$

(Same as $(0,k,1)$ with $k = c$)

.

.

.

Event $(0,c^+,c)$

(Same as $(0,k,k)$ with $k = c$)

Event $(1,0,\cdot)$

$t_{bd} = t_s = x$

with probability $p_1\Phi_o(x)$

Table 4. Formulation of Output Distribution for M|GI|c (Continued)

$$p_1h(t_{bd}) = p_1\beta(t_{bd})\Phi_o(t_{bd})$$

<u>Event (1,1,·)</u>

$$t_{bd} = t_s = t_{ba_1} + x$$

with probability $p_1\Phi_o(x)\Theta_o(x)$

$$p_1h(t_{bd}) = p_1\beta(t_{bd}) \int_o^{t_{bd}} g(t_{bd} - x)\Phi_o(x)\Theta_o(x)dx$$

·

·

·

<u>Event (1,k,·)</u>

$$t_{bd} = t_s = \sum_{j=1}^{k} t_{ba_j} + x$$

with probability $p_1[\prod_{i=2}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j} + x)]\Theta_o(x)\Phi_o(x)$

$$p_1h(t_{bd}) = \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int_o^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Z)dxdt_{ba_k} \cdots dt_{ba_2}$$

Table 4. Formulation of Output Distribution for M|GI|c (Continued)

---

where

$$Z = p_1\beta(t_{bd})[g(t_{bd}-\sum_{j=2}^{k} t_{ba_j} - x)][\prod_{j=2}^{k} g(t_{ba_j})]$$

$$\cdot [\prod_{i=2}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j} + x)]\Theta_o(x)\Phi_o(x)$$

.

.

.

Event $(1,(c-1)^+,\cdot)$

(Same as $(1,k,\cdot)$ with $k = c - 1$)

Events $(1,1,1)$ - $(1,(c-1)^+,c-1)$

(Same as $(0,1,1)$ - $(0,c-1,c-1)$ with p_o replaced by $p_1\gamma_o(t_{bd})$)

.

.

.

Event $(n,0,\cdot)$

$$t_{bd} = t_s = x$$

with probability $p_o[\gamma_o(t_{bd})]^{n-1}\Phi_o(x)$

$$p_n h(t_{bd}) = p_n\beta(t_{bd})[\gamma_o(t_{bd})]^{n-1}\Phi_o(t_{bd})$$

Table 4. Formulation of Output Distribution for M|GI|c (Continued)

Event (n,1,·)

$$t_{bd} = t_{s.} = t_{ba_1} + x$$

with probability $p_n[\gamma_o(t_{bd})]^{n-1}\Theta_o(x)\Phi_o(x)$

$$p_n h(t_{bd}) = p_n \beta(t_{bd})[\gamma_o(t_{bd})]^{n-1} \int_o^{t_{bd}} g(t_{bd}-x)\Theta_o(x)\Phi_o(x)dx$$

·

·

·

Event (n,k,·)

$$t_{bd} = t_{s.} = \sum_{j=1}^{k} t_{ba_j} + x$$

with probability $p_n[\gamma_o(t_{bd})]^{n-1}[\prod_{i=2}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j} + x)]\Theta_o(x)\Phi_o(x)$

$$p_n h(t_{bd}) = \int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int_o^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Z)dxdt_{ba_k} \cdots dt_{ba_2}$$

where $$Z = p_n[\gamma_o(t_{bd})]^{n-1}[g(t_{bd}-\sum_{j=2}^{k} t_{ba_j} - x)][\prod_{j=2}^{k} g(t_{ba_j})]$$

$$\cdot\beta(t_{bd})[\prod_{i=2}^{k} \Theta_o(\sum_{j=i}^{k} t_{ba_j} + x)]\Theta_o(x)\Phi_o(x)$$

Table 4. Formulation of Output Distribution for M|GI|c (Continued)

.

.

.

Event $(n,(c-n)^+,\cdot)$

(Same as $(n,k,\cdot)$ with $k = c - n$)

Events $(n,1,1)$ - $(n,(c-n)^+,c-n)$

(Same as $(0,1,1)$ - $(0,c-n,c-n)$ with p_o replaced by $p_n[\gamma_o(t_{bd})]^n$)

.

.

.

Event $(c-1,0,\cdot)$

$$t_{bd} = t_s = x$$

with probability $p_{c-1}[\gamma_o(t_{bd})]^{c-2}\Phi_o(x)$

$$p_{c-1}h(t_{bd}) = p_{c-1}[\gamma_o(t_{bd})]^{c-2}\beta(t_{bd})\Phi_o(t_{bd})$$

Event $(c-1,1^+,\cdot)$

$$t_{bd} = t_s = t_{ba_1} + x$$

with probability $p_{c-1}[\gamma_o(t_{bd})]^{c-2}\theta_o(x)\Phi_o(x)$

Table 4. Formulation of Output Distribution for M|GI|c (Continued)

$$P_{c-1}h(t_{bd}) = P_{c-1}[\gamma_o(t_{bd})]^{c-2}\beta(t_{bd}) \int_o^{t_{bd}} g(t_{bd}-x)\Theta_o(x)\Phi_o(x)dx$$

<u>Event $(c-1,1^+,1)$</u>

$$t_{bd} = t_{ba_1} + t_{s_1} = t_{ba_1} + x$$

with probability $P_{c-1}[\gamma_o(t_{bd})]^{c-1}\Phi_o(x)$

$$P_{c-1}h(t_{bd}) = P_{c-1}[\gamma_o(t_{bd})]^{c-1} \int_o^{t_{bd}} g(t_{bd}-x)b(x)\Phi_o(x)dx$$

<u>Event $(c^+,0^+,\cdot)$</u> (Unit which replaces departee completes service first.)

$$t_{bd} = t_{s_\cdot} = x$$

with probability $P_{\geq c}[\gamma_o(t_{bd})]^{c-1}$

$$P_{\geq c}h(t_{bd}) = P_{\geq c}[\gamma_o(t_{bd})]^{c-1}\, b(t_{bd})$$

<u>Event $(c^+,0^+,\cdot)'$</u> (Some unit other than the one which replaces departee completes service first.)

$$t_{bd} = t_{s_\cdot} = x$$

Table 4. Formulation of Output Distribution for M|GI|c (Continued)

with probability $P_{\geq c}[\gamma_o(t_{bd})]^{c-2}\Theta_o(t_{bd})$
$P_{\geq c}h(t_{bd}) = P_{\geq c}[\gamma_o(t_{bd})]^{c-2}\Theta_o(t_{bd})\beta(t_{bd})$

of Equation (110) by using another (judiciously chosen) order of summation. Equation (110) represents a departure from the form of previous results in that it is written entirely in the time domain. Possible merits of a Laplace transform domain solution cannot be predicted without a more specific knowledge of the service distributions; however, it may be that some terms can be more easily treated in this domain. In particular, note that each integral can be factored into a coefficient (in t_{bd}) times a convolution. The convolution's transform is, of course, just the product of the transforms of the convolved terms.

Output of GI|GI|c

An output distribution can be formulated for GI|GI|c in much the same way as for M|GI|c. Here, of course, we need expressions for the arrival probabilities when observations are started at some arbitrary time $t = 0$. Let $a(t)$ denote the interarrival time density function; $\nu_o(t)$, the complementary cumulative distribution function for $a(t)$; $\alpha(t)$, the density function for waiting time until the next arrival when observations are started at some arbitrary $t = 0$; and $\eta_o(t)$, the probability that waiting time until the next arrival exceeds t when observations are started at some arbitrary $t = 0$.

Then, by analogy with Equations (106, 107, 109), we have

$$\eta_o(t) = \frac{1}{1/\lambda} \int_t^\infty \nu_o(t) \; , \tag{113}$$

$$\nu_o(t) = \int_t^\infty a(\tau)d\tau \; , \tag{114}$$

and

$$\alpha(t) = - \frac{d\eta_o(t)}{dt} \; . \tag{115}$$

Two changes in Table 4 will suffice to adapt it to the GI|GI|c model. First, g() must be replaced by α() for a first arrival and by a() otherwise. Second, $\Phi_o()$ must be replaced by $\eta_o()$ for Events (n,0,•) and by $\nu_o()$ otherwise. Then, the density function for the time between departures for GI|GI|c becomes

$$h(t_{bd}) = \sum_{n=0}^{c-1} \{p_n[\gamma_o(t_{bd})]^n \int_o^{t_{bd}} \alpha(t_{bd}-x)b(x)\nu_o(x)dx\} \tag{116}$$

$$= \sum_{n=0}^{c-2} \sum_{k=2}^{c-n} \sum_{m=1}^{k} \{\int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int_o^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Y_1)dxdt_{ba_k}\cdots dt_{ba_2}\}$$

$$= \sum_{n=1}^{c-1} \{p_n\beta(t_{bd})[\gamma_o(t_{bd})]^{n-1}\eta_o(x)\}$$

$$+ \sum_{n=1}^{c-1} \{p_n\beta(t_{bd})[\gamma_o(t_{bd})]^{n-1} \int_o^{t_{bd}} \alpha(t_{bd}-x)\theta_o(x)\nu_o(x)dx\}$$

$$+ \sum_{n=1}^{c-2} \sum_{k=2}^{c-n} \{\int_o^{t_{bd}} \int_o^{t_{bd}-t_{ba_2}} \cdots \int^{t_{bd}-\sum_{j=2}^{k} t_{ba_j}} (Y_2)dxdt_{ba_k}\cdots dt_{ba_2}\}$$

$$+ P_{\geq c}[\gamma_o(t_{bd})]^{c-1}b(t_{bd}) + P_{\geq c}[\gamma_o(t_{bd})]^{c-2}\theta_o(t_{bd})\beta(t_{bd}) \ ,$$

where

$$Y_1 = p_n[\gamma_o(t_{bd})]^n[\alpha(t_{bd}-\sum_{j=2}^{k} t_{ba_j}-x)][\prod_{j=2}^{k} a(t_{ba_j})] \tag{117}$$

$$\cdot [\delta_{k,m}b(x) + (1-\delta_{k,m})b(\sum_{j=m+1}^{k} t_{ba_j} + x)]$$

$$\cdot\left[\prod_{\substack{i=2\\ i\neq m+1}}^{k} \Theta_o\left(\sum_{j=i}^{k} t_{ba_j} + x\right)\right]\left[\delta_{k,m} + (1-\delta_{k,m})\Theta_o(x)\right]\nu_o(x) \ ,$$

$$Y_2 = p_n[\gamma_o(t_{bd})]^{n-1}\beta(t_{bd})\left[\alpha\left(t_{bd}-\sum_{j=2}^{k} t_{ba_j} - x\right)\right]\left[\prod_{j=2}^{k} a(t_{ba_j})\right] \quad (118)$$

$$\cdot\left[\prod_{i=2}^{k} \Theta_o\left(\sum_{j=i}^{k} t_{ba_j} + x\right)\right]\Theta_o(x)\nu_o(x) \ ,$$

and $\delta_{k,m}$ is the Kronecker delta of Equation (84).

Output of GI|GI|c with Multiple Inputs

Continuing our analogy between GI-type arrival and service distributions, it is evident that our independence assumptions will also allow treatment of multi-input GI|GI|c models. Suppose that a GI|GI|c queue is fed by ℓ identical, but independent, input sources, each of type GI. Then, for example, the probability that a first arrival occurs at time t after observations are started at some arbitrary t = 0 is

$$\alpha(t)[\eta_o(t)]^{\ell-1} \ . \quad (119)$$

Other combinations can be formed to give the probabilities of various arrival behaviors during the interdeparture interval.

Output of GI|GI|c with Heterogeneous Inputs and Servers

Thus far, we have required that our models be of the strict, homogeneous multichannel types. However, this restriction could be dropped in the models of the present chapter. The condition of independently operating servers (input sources) has made it possible to write the probability of given system behavior as the product of the independent server (input source) behaviors. An implication is that the server (interarrival) time distributions of the various servers (input sources) need not be identical. Of course, many more inter-departure events and their corresponding probabilities must be considered to account for a specific server's (input source's) obtaining (providing) an arrival, etc.

CHAPTER V

DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

Comparisons with Earlier Papers

Burke [8]

The present method for the determination of Poisson output for the $M|M|c$ model is much more lengthy and involved than Burke's differential difference equation approach. However, it does present several dividends:

1. The method is applicable to a great many other $M_n|M_n|1$ models.

2. The initial formulation of possible interdeparture events is in a form which lends itself to analysis of processes with partial information and transient processes.

3. Variations of the method can be used to study more general (e.g. $GI|GI|c$) models.

Finch [25,27]

Equation (86) for the output of a truncated $M|M|c$ queue may be added to Finch's information [25] on the effect of the size of the waiting room. We cannot generally expect the output of a $M_n|M_n|1$ queue to be the same as the input, even if the service time distribution is negative exponential. Finch's result [27], that toleration of an infinite queue and negative exponential servicing are necessary and sufficient conditions for Poisson input to result in Poisson output, does not apply to those ($M_n|M|c$) with state dependent input. (See Reich, below.)

Reich [65]

Reich's partial converse to Burke's theorem [8]; that, for a single channel queue ($M|GI|1$), Poisson input and output imply either negative exponential serving or an impulse function at zero; cannot be extended in at least one direction. Equation (68) implies a partial converse to Reich's partial converse for the case of state dependent input: Poisson output will result for any $M_n|M|c$ queue which possesses inputs of the form $\lambda_n = \lambda$ for $n < c$ and satisfy

$$\sum_{n=c}^{\infty} \prod_{j=c}^{n} \left[\frac{\lambda_n}{(c\mu)^{n-c+1}}\right] = \frac{\lambda}{c\mu-\lambda}$$

and the non-saturation condition of Equation (35) for $n \geq c$. An obvious example is $\lambda_c = c\mu\lambda/(c\mu-\lambda)$ and $\lambda_n = 0$ for $n > c$.

Chang [10]

Chang's result for $GI|GI|1$ requires several difficult computations. In some cases, the present method (which also applies to $GI|GI|c$) may give quicker solutions.

Additional Comments

The tabular or algebraic forms of the output distributions obtained are sufficiently complex as to provide difficulties in application. Possible aids to formulation and resolution of terms were suggested at the times these tables were presented and will not be repeated here. In the event that approximate results are suitable to the intended application, computations might be reduced by making standard numerical approximations. A computer might be programmed to calculate individual terms when the arrival and service parameters are known. Further, calculations might be avoided altogether through some suitable simulation technique (e.g., Dunn, et al., [21]).

Recommendations

In the Introduction, it was remarked that, with the present state of queueing theory, a knowledge of the output distributions at each queue was the missing key to analysis of networks of queues. Earlier papers and the present work represent only a beginning toward comprehensive knowledge of queueing output behavior. For the important applications to queueing network analysis, both simplifications and extensions are needed. The techniques presented by Beneš [5], Chang [10], Conolly [12], Cox [15], Kendall [43], and Moore [59] might be useful in answering both needs.

Extentions are needed to models with more general input and service time distributions (e.g. $G|G|c$) and to queues with batched arrivals and/or departures, priority service, reneging, or other refinements. There are a few papers which might prove useful in extending the present work to such models. For example, Foster [28] and Foster and Perera [29] have given some promising partial information for batched departure epochs in an $E_k|G|1$ queue and Morimura [58] has given some system parameters at a departure epoch for $GI|G|c$ queues.

LITERATURE CITED

1. Akaike, Hirotugu, "On Ergodic Property of a Tandem Type Queueing Process," *Annals of the Institute of Statistical Mathematics* (Japan), Vol. 9 (1957), pp. 13-21.

2. Ancker, C. J., Jr. and A. V. Gafarian, "Queueing with Multiple Poisson Inputs and Exponential Service Times," *Operations Research,* Vol. 9 (1961), pp. 321-327.

3. Ancker, C. J., Jr. and A. V. Gafarian, "Queueing with Reneging and Multiple Heterogeneous Servers," *Naval Research Logistics Quarterly,* Vol. 10 (1963), pp. 125-149.

4. Beightler, C. S., L. G. Mitten, and G. L. Nemhauser, "A Short Table of z-Transforms and Generating Functions," *Operations Research,* Vol. 9 (1961), pp. 574-578.

5. Beneš, V. E., "Combinatory Methods and Stochastic Kolmogorov Equations in the Theory of Queues with One Server," *Transactions of the American Mathematical Society,* Vol. 94 (1960), pp. 282-294.

6. Benson, F. and G. Gregory, "Closed Queueing Systems: A Generalization of the Machine Interference Model," *Journal of the Royal Statistical Society, Series B,* Vol. 23 (1961), pp. 385-393.

7. Boldyref, A., "Determination of the Maximal Steady State Flow of Traffic through a Railroad Network," *Operations Research,* Vol. 3 (1955), pp. 443-465.

8. Burke, Paul J., "The Output of a Queueing System," *Operations Research,* Vol. 4 (1956), pp. 699-704.

9. Burke, P. J., "The Dependence of Delays in Tandem Queues," *Annals of Mathematical Statistics,* Vol. 35 (1964), pp. 874-875.

10. Chang, Wei, "Output Distribution of a Single-Channel Queue," *Operations Research,* Vol. 11 (1963), pp. 620-623.

11. Cohen, J. W., "On the Queueing Processes of Lanes, *Philips Technical Report,* 1956.

12. Conolly, B. W., "A Difference Equation Technique Applied to the Simple Queue," *Journal of the Royal Statistical Society, Series B,* Vol. 20 (1958), pp. 165-167.

13. Conolly, B. W., "A Difference Equation Technique Applied to the Simple Queue with Arbitrary Interval Distribution," *Journal of the Royal Statistical Society, Series B,* Vol. 20 (1958), pp. 167-175.

14. Conolly, B. W. "Queueing at a Single Serving Point with Group Arrival," *Journal of the Royal Statistical Society, Series B,* Vol. 22 (1960), pp. 258-298.

15. Cox, D. R., "Some Statistical Methods Connected with Series of Events," *Journal of the Royal Statistical Society, Series B,* Vol. 17 (1955), pp. 129-164.

16. Cox, D. R. and W. L. Smith, *Queues,* Methuen's Monographs on Statistical Subjects, Methuen and Co., Ltd., London; John Wiley and Sons, Inc., New York; 1961, xii + 180 pp.

17. Daru, E., "Mèthodes de Resolution pour Quelques Problèmes de Files d'Attente Comportant des 'Serveurs' d'Efficacités Différentes," *Revue Française de Recherche Opérationnelle,* Vol. 2 (1958), pp. 137-152.

18. DeBaun, R. M. and S. Katz, "An Approximation to Distributions of Summed Waiting Times," *Operations Research,* Vol. 7 (1959), pp. 811-813.

19. Descamps, René, "Calcul Direct des Probabilités d'Attente dans une File," *Revue Française de Recherche Opérationnelle,* Vol. 3 (1959), pp. 88-100.

20. Doig, A. "A Bibliography on the Theory of Queues," *Biometrika,* Vol. 44 (1957), pp. 490-514.

21. Dunn, Paul F., C. D. Flagle, and P. A. Hicks, "The Queuiac: An Electromechanical Analog for the Simulation of Waiting-Line Problems," *Operations Research,* Vol. 4 (1956), pp. 648-662.

22. Fagen, R. E. and J. Riordan, "Queueing Systems for Single and Multiple Operations," *Journal of the Society for Industrial and Applied Mathematics,* Vol. 3 (1955), pp. 73-79.

23. Feller, Wm., "On Boundary Conditions for the Kolmogorov Differential Equations," *Annals of Mathematics,* Vol. 65 (1957), pp. 527-570.

24. Feller, Wm., *An Introduction to Probability Theory and Its Applications,* John Wiley and Sons, Inc., New York, 1950, xv + 461 pp.

25. Finch, P. D., "The Effect of the Size of the Waiting Room on a Simple Queue," *Journal of the Royal Statistical Society, Series B,* Vol. 20 (1958), pp. 182-186.

26. Finch, P. D., "Cyclic Queues with Feedback," *Journal of the Royal Statistical Society, Series B,* Vol. 21 (1959), pp. 153-157.

27. Finch, P. D., "The Output Process of the Queueing System M|G|1," *Journal of the Royal Statistical Society, Series B,* Vol. 21 (1959), pp. 375-380.

28. Foster, F. G., "Batched Queueing Processes," *Operations Research,* Vol. 12 (1964), pp. 441-449.

29. Foster, F. G. and A. G. A. D. Perera, "Queues with Batch Departures, II," *Annals of Mathematical Statistics,* Vol. 35 (1964), pp. 1147-1156.

30. Gani, J., "First Emptyness of Two Dams in Parallel," *Annals of Mathematical Statistics,* Vol. 32 (1961), pp. 219-229.

31. Ghosal, A., "Queues in Series," *Journal of the Royal Statistical Society, Series B,* Vol. 24 (1962), pp. 359-363.

32. Good, I. J., "The Number of Individuals in a Cascade Process," *Proceedings of the Cambridge Philosophical Society,* Vol. 45 (1948), pp. 360-363.

33. Gumbel, Harold, "Waiting Lines with Heterogeneous Servers," *Operations Research,* Vol. 8 (1960), pp. 504-511.

34. Haight, Frank A., "Two Queues in Parallel," *Biometrika,* Vol. 45 (1958), pp. 401-410.

35. Heathcote, C. R., "The Time-Dependent Problem for a Queue with Preemptive Priorities," *Operations Research,* Vol. 7 (1959), pp. 670-680.

36. Hunt, G. C., "Sequential Arrays of Waiting Lines," *Operations Research,* Vol. 4 (1956), pp. 674-683.

37. Jackson, J. R., "Networks of Waiting Lines," *Operations Research,* Vol. 5 (1957), pp. 518-521.

38. Jackson, J. R., "Multiple Servers with Limited Waiting Space," *Naval Research Logistics Quarterly,* Vol. 5 (1958), pp. 315-321.

39. Jackson, R. R. P., "Queueing Systems with Phase Type Service," *Operations Research Quarterly,* Vol. 5 (1954), pp. 109-120.

40. Jackson, R. R. P., "Random Queueing Processes with Phase-Type Service," *Journal of the Royal Statistical Society, Series B,* Vol. 18 (1956), pp. 129-132.

41. Jewell, William S., "Multiple Entries in Traffic," *Journal of the Society for Industrial and Applied Mathematics,* Vol. 11 (1963), pp. 872-885.

42. Karlin, S., R. C. Miller, Jr., and N. V. Prabhu, "Note on a Moving Single Server Problem," *Annals of Mathematical Statistics,* Vol. 30 (1959), pp. 243-246.

43. Kendall, D. G., "Stochastic Processes Occurring in the Theory of Queues and Their Analysis by the Method of the Imbedded Markov Chain," *Annals of Mathematical Statistics,* Vol. 24 (1953), pp. 338-354.

44. Kingman, J. F. C., "Two Similar Queues in Parallel," *Annals of Mathematical Statistics,* Vol. 32 (1961), pp. 1314-1323.

45. Koenigsberg, E., "Cyclic Queues," *Operational Research Quarterly,* Vol. 9 (1958), pp. 22-35.

46. Koenigsberg, E., "Finite Queues and Cyclic Queues," *Operations Research,* Vol. 8 (1960), pp. 246-253.

47. Krishnamoorthi, B., "On Poisson Queue with Two Heterogeneous Servers," *Operations Research,* Vol. 11 (1963), pp. 321-330.

48. Lindley, D. V., "The Theory of Queues with a Single Server," *Proceedings of the Cambridge Philosophical Society,* Vol. 48 (1952), pp. 277-289.

49. Loynes, R. M., "The Stability of a System of Queues in Series," *Proceedings of the Cambridge Philosophical Society,* Vol. 58 (1962), pp. 497-500.

50. Luchak, George, "The Continuous Time Solution of the Equations of the Single Channel Queue with a General Class of Service-Time Distributions by the Method of Generating Functions," *Journal of the Royal Statistical Society, Series B,* Vol. 20 (1958), pp. 176-181.

51. Lunger, G. F., "Bibliography on Queueing Theory," Sperry Rand Corporation, Univac Division, 1959.

52. Makino, Toji, "Tandem-Gate Machi-Gyōretsu Ni Tsuiteno Ni San No Mondai," ("Some Problems on Tandem Queues"), *Management Science* (Japan), Vol. 6 (1963), pp. 246-257.

53. Makino, Toji, "On the Mean Passage Time Concerning Some Queueing Problems of the Tandem Type," *Journal of the Operations Research Society of Japan,* Vol. 7 (1964), pp. 17-47.

54. Marshal, A. W. and E. Reich, "Study of the Characteristics of Queues in Tandem," *Operations Research,* Vol. 4 (1956), p. 386.

55. Masterson, Gregory E. and Seymour Sherman, "On Queues in Tandem," *Annals of Mathematical Statistics,* Vol. 34 (1963), pp. 300-307.

56. McMillan, B. and J. Riordan, "A Moving Single Server Problem," *Annals Mathematical Statistics,* Vol. 28 (1957), pp. 471-478.

57. Molina, E. C., *Poisson's Exponential Binomial Limit,* D. Van Nostrand Company, Inc., New York, 1949, 47 pp.

58. Morimura, Hidenori, "On the Relation Between the Distributions of the Queue Size and the Waiting Time," *Kodia Mathematical Seminar Reports* (Japan), Vol. 14 (1962), pp. 6-19.

59. Moore, P. G., "Interval Analysis and the Logarithmic Transformation," *Journal of the Royal Statistical Society, Series B.,* Vol. 20 (1958), pp. 187-192.

60. Morris, W. T., *Analysis for Materials Handling Management,* Richard D. Irwin, Inc., Homewood, Illinois, 1962, ix + 204 pp.

61. Nelson, Ross T., "Waiting-Time Distributions for Application to a Series of Service Centers," *Operations Research,* Vol. 6 (1958), pp. 856-862.

62. O'Brian, G. G., "The Solution of Some Queueing Problems," *Journal of the Society for Industrial and Applied Mathematics,* Vol. 2 (1954), pp. 133-142.

63. Phillips, C. R., Jr., "A Variable-Channel Queueing Model with a Limited Number of Channels," Thesis, Georgia Institute of Technology, 1960, 77 pp.

64. Reich, E., "Birth-Death Process and Tandem Queues," The Rand Corporation, Paper P863, 1956.

65. Reich, E., "Waiting Times When Queues Are in Tandem, *Annals of Mathematical Statistics,* Vol. 28 (1957), pp. 768-773.

66. Reich, E., "Note on Queues in Tandem," *Annals of Mathematical Statistics,* Vol. 34 (1963), pp. 338-341.

67. Romani, J. "Distribution of the Algebraic Sum of Poisson Variables," *Trabajos de Estadistica,* Vol. 7 (1956), pp. 175-181.

68. Romani, J. "Un Modelo de la Teoria de Colas con Numero Variable de Canales," *Trabajos de Estadistica,* Vol. 8 (1957), pp. 175-189.

69. Saaty, Thomas L., *Elements of Queueing Theory with Applications,* McGraw-Hill Book Co., Inc., New York-Toronto-London, 1961, xv + 423 pp.

70. Sacks, J., "Ergodicity of Queues in Series," *Annals of Mathematical Statistics*, Vol. 31 (1960), pp. 579-588.

71. Sasieni, M., A. Yaspan, and L. Friedman, *Operations Research: Methods and Problems*, John Wiley and Sons, Inc., New York-London, 1959, pp. 125-154.

72. Stephan, F. F., "Two Queues with a Priority Rule and a Single Queue with Interrupted Service," *Operations Research*, Vol. 4 (1956), p. 386.

73. Stephan, F. F., "Two Queues under Preemptive Priority with Poisson Arrival and Service Rates," *Operations Research*, Vol. 6 (1958), pp. 399-418.

*74. Suzuki, Takeji, "Two Queues in Series," *Journal of the Operations Research Society of Japan*, Vol. 5 (1962/1963), pp. 125-154.

75. Takács, L., "Investigation of Waiting Time Problems by Reduction to Markov Processes," *Acta Mathematica Academia Scientiarum Hungaricae* (Hungary), Vol. 6 (1955), pp. 101-129.

76. Takács, L., "A Single-Server Queue with Feedback," *Bell System Technical Journal*, Vol. 42 (1963), pp. 505-519.

77. Takács, L., *Introduction to the Theory of Queues*, University Texts in the Mathematical Sciences, Oxford University Press, New York, 1962, x + 268 pp.

78. Tanner, J. C., "A Problem of Interference between Two Queues," *Biometrika*, Vol. 40 (1953), pp. 58-69.

79. Ventura, E., "Sur l'Utilization des Intégrales de Contour dans les Problèmes de Stocks et de Délais d'Attente, *Management Science*, Vol. 6 (1960), pp. 423-443.

80. White, H. C. and L. S. Christe, "Queueing with Preemptive Priorities or with Breakdown," *Operations Research*, Vol. 6 (1958), pp. 79-95.

81. Wilkins, Coleridge A., "On Two Queues in Parallel," *Biometrika*, Vol. 47 (1960), pp. 198-199.

---

*According to L. Takács (*Mathematical Reviews*, Vol. 29 (1965), #2873, p. 559), Suzuki's results are not correct ". . . because his proof is based on the false assumption that the queue sizes immediately before arrivals in the second queue form a Markov chain."